NONRECURSIVE DIGITAL IMAGE RESTORATION

A thesis presented for the degree of
Doctor of Philosophy in Electrical Engineering
in the University of Canterbury,
Christchurch, New Zealand.

by

M.J. McDONNELL B.Sc. (Hons)

1975

ABSTRACT

The problem of digitally restoring blurred images is considered. The subject of image restoration is reviewed in detail and a comprehensive notation for image restoration is developed. Degraded images are divided into classes $\mathcal{G}$ and $\mathcal{I}$ - those which are and are not respectively truncated by their recording frames. It is shown that conventional restoration techniques only work well for images of class $\mathcal{I}$. Restoration techniques for dealing with class $\mathcal{G}$ images are presented, as well as improved techniques for dealing with class $\mathcal{I}$ images. These new techniques are shown to be both efficient and practical through the use of computer simulations, and through the restoration of optically degraded images.

The problem of designing an optimum nonrecursive digital filter array of a given number of elements is solved in both one and two dimensions. It is demonstrated that this new design method is superior to previous techniques.

A sampling theory appropriate for deconvolution problems is presented. New sampling functions are introduced, in both one and two dimensions, which overcome the "picket fence" effect associated with the sinc sampling function. The concept of a "line-segment-limited" function is defined. It is shown that a pseudo noise level is introduced by the approximation that the point spread function is line-segment-limited. It is this pseudo noise

level, and not aliasing errors, that restricts the choice of a sampling rate less than the Nyquist rate. New criteria are presented which allow the choice of a sampling rate considerably less than the Nyquist rate.

ACKNOWLEDGEMENTS

I am deeply indebted to my supervisor, Professor R.H.T. Bates, for his enthusiastic guidance and encouragement, and also for many useful discussions during the course of this project. I also wish to thank Dr T.M. Peters who supervised me for some months while Professor Bates was on sabbatical leave.

I am grateful to my colleagues Dr P.T. Gough, Mr D.J.N. Wall and Mr R.M. Lewitt in the Electrical Engineering Department for many stimulating discussions on aspects of this work. I especially wish to thank Mr W.K. Kennedy for his valuable assistance.

The assistance received from a number of organisations is gratefully acknowledged. In particular I wish to thank Detective Superintendent E.G. Perry and Detective Inspector J.W. Wooders of the Christchurch Police District, Mr G.R. Scott of the Chemistry Division DSIR in Christchurch, and also the Medical Physics Department of the Christchurch Hospital for use of the electrostatic lineprinter.

I wish to thank my wife Elizabeth for her patience and encouragement.

The financial support of a postgraduate scholarship from the University Grants Committee is gratefully acknowledged.

TABLE OF CONTENTS

Page

Page

Page

GLOSSARY

Listed below are those symbols used in this thesis which are not always defined in the immediate context in which they are used.

FFT	fast Fourier transform
FIR	finite impulse response
FT	Fourier transform
MSE	mean-square error
MTF	modulation transfer function
psf	point spread function
PSWF	prolate spheroidal wave function
SI	spatially invariant
SV	spatially varying
1-D	one dimension or one-dimensional
2-D	two dimensions or two-dimensional
A	extent of Γ in x-direction
A'	extent of Γ in y-direction
$\mathscr{A}(x,y)$	available degraded image
$b(x,y)$	ideal degraded image
$b(r;\theta)$	$b(x,y)$ in polar coordinates
$b_m(r)$	the mth term in the angular Fourier series expansion of $b(r;\theta)$
$b_{m,n}$	the nth coefficient in the Fourier-Bessel series expansion of $b_m(r)$
$\hat{b}(x,y)$	an estimate of $b(x,y)$
$B(u,v)$	FT of $b(x,y)$
$B(\rho;\phi)$	FT of $b(r;\theta)$

$B_m(\rho)$	FT of $b_m(r)$
$\hat{B}(u,v)$	FT of $\hat{b}(x,y)$
$c(x,y)$	contamination in $f(x,y)$
$c_q(x)$	the part of $c(x,y)$ in 1-D which is not removable
$c_r(x)$	the part of $c(x,y)$ in 1-D which is removable
$\mathrm{comb}(x)$	$= \sum_{n=-\infty}^{\infty} \delta(x-n)$
$C(u,v)$	FT of $c(x,y)$
$C_q(u)$	FT of $c_q(x)$
$C_r(u)$	FT of $c_r(x)$
C_n	Noise amplification factor
C_n^{max}	maximum acceptable noise amplification factor
$\mathscr{D}\{p\}$	preliminary degradation of p
e	exponential constant
$e(x,y)$	ideal overlapped degraded image
$e'(x)$	defined by (5.4)
$\hat{e}(x)$	an approximation to $e(x,y)$ in 1-D
$\hat{e}^i(x)$	initial estimate of $\hat{e}(x)$
$E(u,v)$	FT of $e(x,y)$
E_m	mth sample of $E(u)$
$\hat{E}_m$	mth sample of $\hat{E}(u)$
$E_{uB}(u)$	upper bound on $\lvert E(u)\rvert$
E_{uBm}	mth sample of $E_{uB}(u)$
$f(x,y)$	final degraded image
f_j	jth sample of $f(x,y)$ in 1-D

$f_{k,l}$	sample of f(x,y)
$\underset{\sim}{f}$	column vector (f_j)
$\hat{f}(x)$	estimate of f(x,y) in 1-D
$f_D(x)$	decontaminated blurred image in 1-D
$f_B(x,y)$	inverse FT of $F_B(u,v)$
F(u,v)	FT of f(x,y)
$F_{m,n}$	sample of F(u,v)
F_k	sample of F(u,v) in 1-D
$F_B(u,v)$	result of bandlimiting F(u,v)
$\mathscr{F}$	Fourier space
G,G'	real variables
$\mathscr{G}$	the general class of degraded images which are truncated by their recording frames
h(x,y)	SI psf
$h(x,y,\alpha,\beta)$	SV psf
h_k	sample of h(x,y) in 1-D
h(r)	radially symmetric psf
$h_{j,k}$	sample of h(x,y)
$h_B(x,y)$	inverse FT of $H_B(u,v)$
$h_c(x,y)$	the line-segment-limited psf
$h_s(x,y)$	the sampled psf
$\underset{\sim}{h}$	column vector (h_j)
$\bar{h}(x,y)$	inverse psf (inverse FT of $\bar{H}(u,v)$)
H(u,v)	FT of h(x,y)
$H(\rho)$	FT of h(r)

$\overline{H}(u,v)$	inverse filter
$H_B(u,v)$	result of bandlimiting $H(u,v)$
H_m	mth sample of $H(u)$
i	$= \sqrt{-1}$
I	set of integers
I_+	set of positive integers
$\mathscr{I}$	image space
j,k,l,m,n,j_m,k_m	integer variables
$j_{m,n}$	nth positive zero of $J_m(\cdot)$
$J_m(\cdot)$	mth order Bessel function of the first kind
J,K,M,N,M_1,M_2,N_1,N_2	integer constants
ℓ	real constant defined in (3.6)
L	extent of T in x-direction
L'	extent of T in y-direction
$\mathscr{L}\{\cdot\}$	frame function defined by (2.8)
$n(x,y)$	noise in $f(x,y)$
n_j	jth sample of $n(x,y)$ in 1-D
n'_j	jth sample of noise after filtering
N_f	Nyquist frequency corresponding to $f(x,y)$
N_z	number of positive zeros of $H(\rho)$ under consideration
$p(x,y)$	original image
$p(r;\theta)$	$p(x,y)$ in polar coordinates
$p_{j,k}$	sample of $p(x,y)$
$\hat{p}(x,y)$	estimate of $p(x,y)$
$\hat{p}_k$	sample of $\hat{p}(x,y)$ in 1-D
$P(u,v)$	FT of $p(x,y)$

$P(\rho;\phi)$	FT of $p(r;\theta)$
$P_B(u,v)$	result of bandlimiting $P(u,v)$
P_m	mth sample of $P(u)$
$P_m(\rho)$	mth term in truncated angular Fourier series expansion of $P(\rho;\phi)$
P_{uBm}	an upper bound on $\lvert P_m \rvert$
$\hat{P}(u,v)$	FT of $\hat{p}(x,y)$
r	polar radial coordinate in $\mathscr{I}$
$\text{rect}(x)$	$= 1, \quad \lvert x \rvert \leqslant \frac{1}{2}$
	$= 0, \quad \lvert x \rvert > \frac{1}{2}$
R_+	set of positive real numbers
$s(x,y)$	resultant psf
$s_1(x)$	1-D sampling function for line-segment-limited functions
$s_2(x,y)$	2-D sampling function for line-segment-limited functions
$\text{sgn}(x)$	$= -1, \quad x < 0$
	$= 0, \quad x = 0$
	$= 1, \quad x > 0$
$\text{sinc}(x)$	$= \sin(\pi x)/\pi x$
$\text{step}(x)$	$= 0, \; x < 0$
	$= 1, \; x \geqslant 0$
$\mathscr{S}$	special class of degraded images which are not truncated by their recording frames
$t(x,y)$	truncated ideal image
t	as superscript denotes matrix transpose
$T_\alpha(\cdot)$	non-linear operator
u,v	Cartesian coordinates in $\mathscr{F}$

u_c, u_c'	frequency cutoffs
u_m	mth sample of u
w, w_m'	real constants
$w'(x)$	window function defined by (5.34)
w_n	sample of $\underset{\sim}{W}$
$\underset{\sim}{W}$	filter array
$\underset{\sim}{W}_I$	ideal filter array
x, y	Cartesian coordinates in $\mathscr{I}$
x_m	mth sampling position on x-axis
X_m, x_m', y_m, Y_m	real constants
$z(x,y)$	amplitude distribution of coherent light incident on front focal plane
$Z(u,v)$	FT of $z(x,y)$
α, β	dummy variables
γ	arbitrary point in $\mathscr{I}$
γ_f	film gamma
Γ	frame containing $f(x,y)$
Γ/Ψ	$= \{x: x \in \Gamma, x \notin \Psi\}$
$\delta(x,y)$	Dirac delta function
Δ	sampling distance
Δ_a	sampling distance adequate to display detail of interest in $p(x,y)$
ζ_m	mth right half plane complex zero of $H(u)$
η_m	complex component of ζ_m
θ	polar angular coordinate in $\mathscr{I}$

Θ	finite domain of $\mathscr{J}$
Λ	frame containing b(x,y)
Ω/Γ	$= \{\gamma: \gamma \in \Omega, \gamma \notin \Gamma\}$
μ	real constant
ξ_m	real part of ζ_m
Ξ	defined by (3.6)
π	real constant
Π	indicates product of terms
ρ	polar radial coordinate in $\mathscr{F}$
σ_h	noise level due to approximation $h = h_c$
σ_n	original noise level on f
σ_n^{max}	maximum acceptable noise level
Σ	indicates summation
Υ	psf frame
$\overline{\Upsilon}$	inverse psf frame
ϕ	polar angular coordinate in $\mathscr{F}$
Φ	$= \|C(x,y)\|/\|P(x,y)\|$
Ψ	frame containing p(x,y) for a class $\mathscr{S}$ image
ω	arbitrary point in $\mathscr{F}$
Ω	frame containing p(x,y)
*	superscript denotes complex conjugate
$\|\cdot\|$	modulus
$\partial/\partial x$	differential with respect to x
$\in$	belongs to (a set)
$\otimes$	denotes convolution
∞	infinity
$\forall$	for all members (of a set)

$\equiv$	is equivalent to
$\subset$	is included in (a set)

PREFACE

During the last two decades, dramatic advances in technology have made practical the widespread use of sophisticated signal processing techniques. This is mainly due to increases in the availability, size and speed of digital computers, and to the development of efficient algorithms such as the fast Fourier transform algorithm (FFT). For example, using the FFT it is possible to do in minutes computations that would otherwise have taken years.

This progress in signal processing techniques has made it feasible to process images in a computer, and the space program has given impetus to research in this area. Image processing has applications not only in aerial photography but also in other areas such as electron microscopy, medical imaging and astronomy. A problem which often occurs in each of the above areas is that an image may be unavoidably blurred in such a way that it is difficult to extract the required information from it visually. In general this information is not lost, but is masked by the blurring. It is therefore of interest to study image restoration which is concerned with removing the blurring.

Image restoration may be performed by a combination of holography and standard photographic and optical techniques. Alternatively, the image can be digitised by scanning it with a computer-controlled microdensitometer, so

that it can be stored as discrete "pixels" (picture elements) in computer memory - the restoration is then carried out in a digital computer. The interest in this thesis is in digital techniques which, in my experience, are superior to optical methods because of their greater versatility and accuracy.

This thesis is concerned with the development of efficient techniques for the digital restoration of blurred images. It is shown that conventional image restoration techniques are unsatisfactory, and new restoration techniques are developed which overcome the deficiencies of conventional methods. New material is included in chapters 2-7.

In chapter 1, the subject of image processing is reviewed in detail. The problem is defined, and methods of deducing the point spread function characterising the blurring are reviewed. Digital and optical image restoration techniques are reviewed separately and finally compared.

In chapter 2, a comprehensive notation for use in image restoration is introduced. In addition, some of the problems associated with conventional restoration techniques are discussed. Degraded images are divided into the classes $\mathcal{G}$ and $\mathcal{S}$ - those that are and are not respectively truncated by their recording frames. Most images are of class $\mathcal{G}$. It is shown that conventional restoration techniques work well for images of class $\mathcal{S}$, but are unsoundly based for images of class $\mathcal{G}$. The circular

property of convolution, which is exhibited when the FFT is used, is also discussed.

A problem which sometimes occurs in the restoration of class $\mathcal{S}$ images is that zeros of the modulation transfer function (MTF) of the blurring may lie on sampling points. In chapter 3, this problem is discussed in general and it is shown in particular how it may be overcome in the case of uniform linear blur.

Two techniques for the restoration of rectangular class $\mathcal{G}$ images are presented in chapter 4. The better of these two methods is illustrated with a number of restorations of optically degraded images. This technique is shown to be practical and very efficient.

In chapter 5, it is shown how the simple but effective methods of chapter 4 can be modified to take advantage of knowledge of the zeros of the MTF. The resulting restoration method is shown to be less efficient but more accurate than the best method of chapter 4. A simple technique is presented for the restoration of class $\mathcal{G}$ images degraded by uniform linear blur. The possibility of restoring class $\mathcal{G}$ images recorded in circular rather than rectangular frames is also considered.

It is possible to restore degraded images by direct convolution with a nonrecursive digital filter array of finite length. This is considered in detail in chapter 6 and an improved technique is described for deriving an optimum finite filter array of a given length.

In chapter 7, a sampling theory appropriate for

deconvolution is presented. New sampling functions are introduced, in both one and two dimensions, which overcome the "picket fence" effect associated with the sinc sampling function. The concept of a "line-segment'limited" function is defined. It is shown that a pseudo noise level is introduced by the approximation that the point spread function is line-segment-limited. It is this pseudo noise level, and not aliasing errors, that restricts the choice of a sampling rate less than the Nyquist rate. New criteria are presented which allow the choice of a sampling rate considerably less than the Nyquist rate.

Chapter 8 consists of concluding remarks and some suggestions for further research and development related to the results presented in this thesis.

Papers completed to date on topics relevant to the material presented in this thesis are as follows:

Stedman, G.E., McDonnell, M.J. 1973 "Optimisation of coherence properties of thermal sources by spatial filtering," J. Opt. Soc. Am. 63, 1222-1224.

Bates, R.H.T., Kennedy, W.K., McDonnell, M.J. 1974 "Efficient digital restoration of images blurred by linear motion," Letters in Applied and Engineering Sciences 2, 133-143.

McDonnell, M.J., Bates, R.H.T. 1975 "Preprocessing of degraded images to augment existing restoration methods," Computer Graphics and Image Processing 4, 25-39.

McDonnell, M.J., Bates, R.H.T. 1975 "Restoring parts of scenes from blurred photographs," Optics Communications 13, 347-349.

McDonnell, M.J. 1975 "Nonrecursive image restoration using a finite filter array," Optik 43, 159-174.

Bates, R.H.T., Napier, P.J., McKinnon, A.E., McDonnell, M.J. 1975 "Self-consistent deconvolution: I: Theory," to be published in Optik.

McKinnon, A.E., McDonnell, M.J., Napier, P.J., Bates, R.H.T. 1975 "Self-consistent deconvolution: II: Applications," to be published in Optik.

McDonnell, M.J. 1975 "A sampling function appropriate for deconvolution," to be published in IEEE Trans. on Information Theory.

McDonnell, M.J., Kennedy, W.K., Bates, R.H.T. 1975 "Identifying and overcoming practical problems of digital image restoration," submitted to the New Zealand Journal of Science.

McDonnell, M.J., Bates, R.H.T. 1975 "Digital restoration of an image of Betelgeuse," submitted to Ap. J.

Bates, R.H.T., McDonnell, M.J., Gough, P.T. 1975 "Processing images of objects viewed through randomly fluctuating media," invited paper submitted to Proc. IEEE.

CHAPTER ONE

REVIEW OF IMAGE RESTORATION

1.1 INTRODUCTION

The subject of image processing is interdisciplinary. It has been defined by Huang, Schreiber and Tretiak (1971) as dealing with the manipulation of data which are inherently two-dimensional (2-D) in nature. It includes image restoration, image enhancement, pictorial pattern recognition and image coding, and also finds application in such diverse areas as medical imaging (Hall et al 1971, Mersereau and Oppenheim 1974), space exploration and aerial reconnaissance (Sawchuk 1973), electron microscopy (Saxton 1974), astronomy (Bates and Gough 1975), geophysics (Backus and Gilbert 1970) and police work.

The interest here is in image restoration, which is closely allied to image enhancement. Some authors (Stroke 1972, Frieden 1973) consider these two subjects to be identical, while others (Huang et al 1971) take image restoration to be part of image enhancement. In general, however, as in this thesis, the two subjects are treated as distinct (Andrews 1974, Billingsley 1973, Rosenfeld 1972). In line with Andrews (1974) the distinction is made as follows. Image enhancement is concerned with improving, for human viewing or subsequent machine processing, the presentation of an image whose essential information is

visually apparent before processing (this improvement may include such aspects as noise reduction, compensation for nonlinearities of the recording medium, contrast adjustment, edge sharpening, etc.). Image restoration is concerned with the reconstruction of an ideal or original image by inversion of some degradation processes. It usually involves the uncovering of information which has been masked by the degradation. Once a degraded image has been restored, it may be desirable to use enhancement techniques to improve its display.

Image restoration techniques are classified as *a posteriori* or *a priori*. In the former, the restoration is carried out after the degradation has occurred. In the latter, the imaging system is designed so as to minimise the effect of an anticipated degradation, or to simplify subsequent application of *a posteriori* techniques. *A priori* methods have found their widest applications in attempts to overcome the effect of atmospheric turbulence on astronomical imaging (Labeyrie 1970, Bates and Gough 1975), and also in compensating for camera motion (NASA/ERC Seminar 1968). Shutter modulation techniques (Bryngdahl and Lohmann 1968) have also been proposed for modifying the degradation. *A priori* methods have usually been designed for particular applications. In this thesis the interest is in the more general *a posteriori* methods.

1.2 THE PROBLEM

In discussing *a posteriori* image restoration, images

are considered to be 2-D and, for simplicity, monochromatic. Such an image is represented as a real function of two coordinates (x,y) representing the intensity of the images at a point in image space $\mathscr{I}$. The ideal image that would be obtained with a perfect recording system is denoted as $p = p(x,y)$. Since any practical recording system is imperfect, the available image is necessarily degraded. It is denoted by $\mathscr{A} = \mathscr{A}(x,y)$. The purpose of image restoration is to obtain from $\mathscr{A}$ an estimate of p. This estimate which is denoted by $\hat{p} = \hat{p}(x,y)$, minimises in some sense - e.g., the mean-square error (MSE) criterion (Sondhi 1972) - the difference between itself and p.

The degrading system could conceivably be very complicated. However, in certain cases which have practical application it has been modelled (Huang et al 1971) by the block diagram shown in Fig. 1.1.

The ideal image p is acted on by a linear system with an impulse response, or point spread function (psf), which is spatially varying (SV). At a general spatial point (α,β) this psf is given by $h = h(x,y,\alpha,\beta)$ so that

$$b = b(x,y) = \int_{-\infty}^{\infty}\int h(x,y,\alpha,\beta)p(\alpha,\beta)d\alpha d\beta. \qquad (1.1)$$

The available degraded image $\mathscr{A}$ is in general a *non*linear function $T_\alpha(\cdot)$ of the sum of b and the noise $n = n(x,y)$, which is usually assumed to be additive (this assumption is discussed in 1.4.4). This gives

$$\mathscr{A} = \mathscr{A}(x,y) = T_\alpha(b(x,y) + n(x,y)). \qquad (1.2)$$

T_α is a characteristic of the detector used to record the intensity of the degraded image. In practice, an estimate of T_α is usually available. It is assumed here, as is most often the case, that the detector is photographic film, so that (1.2) becomes

$$\mathscr{A} = (b+n)^{\gamma_f} \tag{1.3}$$

where γ_f is called the gamma of the film (Goodman 1968). It is possible to estimate γ_f from the characteristics of the film and the developer used on the film, and from the duration of the development time (Goodman 1968).

Harris (1968) has shown that errors in γ_f in (1.3) have only a slight effect on the quality of restoration. Therefore most authors (e.g., Sondhi 1972, Frieden 1974) assume that the gamma can be corrected easily, so that it makes practical sense to define a final recorded image, $f = f(x,y)$, by

$$f = b+n. \tag{1.4}$$

It is the restoration of p from f which is the main concern here.

Usually h is taken to be spatially invariant (SI):

$$h = h(x,y,\alpha,\beta) = h(x-\alpha,y-\beta). \tag{1.5}$$

This is often true in practice. When h is SV it is often possible to define some geometrical or other transformation (Sawchuk 1973, Robbins and Huang 1972) by which the degraded image is transformed to one in which h is SI.

In some cases (Huang et al 1971) h varies only slowly with position, so that the image may be segmented (Granger 1968) and considered to be SI over each segment. Some restoration techniques discussed later allow h to be SV. However, unless stated otherwise, h is assumed in this thesis to be SI. This makes it possible to express (1.4) as

$$f = p \otimes h + n \tag{1.6}$$

where $\otimes$ denotes convolution.

This is now in the form of a general 2-D deconvolution problem. However, because image intensities are being considered, it is also true that

$$f,b,p,h \in R_+ \tag{1.7}$$

where R_+ is the field of non-negative, real numbers. This additional condition is used in "positive" restoration methods which are discussed in 1.4.3.

It is worthwhile noting here that, by including (1.7), attention is restricted to incoherent imaging systems (as is assumed in most of this thesis). However, restoration methods which do not invoke (1.7) are also applicable to coherent imaging systems (e.g., phase contrast or amplitude contrast electron microscopy (Erickson 1973)).

The Fourier transform (FT) of a quantity is identified by the upper case form of the lower case English letter which denotes the quantity, e.g.

$$F = F(u,v) = \int_{-\infty}^{\infty}\int f(x,y)e^{2\pi i(xu+yv)}dxdy, \tag{1.8}$$

where u and v are the Cartesian coordinates in Fourier space $\mathcal{F}$ and $i = \sqrt{-1}$. Coordinates in $\mathcal{F}$ may be interpreted as spatial frequencies. A component at (u,v) in $\mathcal{F}$ corresponds to a sinusoidal variation in $\mathcal{I}$ with frequency components u and v in the x and y directions respectively.

1.3 PSF IDENTIFICATION

In order to solve (1.6) for p it is necessary to know h. Methods for determining h have been reviewed by Huang et al (1971). Sometimes h is known from a combination of mathematical analysis together with calibration measurements of the imaging system (NASA/ERC Seminar 1968, Erickson and Klug 1971) or from knowledge of camera motion during the exposure (NASA/ERC Seminar 1968). The general form of h may, on occasions, be known initially but quantitative details may have to be deduced. In other cases, h may be completely unknown.

1.3.1 Using Zeros of H

Table 1.1 lists the forms assumed by h for a number of common degradations, following Roetling et al (1968). The table also lists the optical transfer function, $H = H(u,v)$, and the positive points in $\mathcal{F}$ at which H is zero for each of the degradations. The cylindrical polar coordinates $(r;\theta)$ and $(\rho;\phi)$ used in Table 1.1 correspond to (x,y) and (u,v) respectively. Throughout this thesis, variables separated by a semi-colon are always taken to be polar coordinates. Also in Table 1.1, rect(x) is defined by

$$\begin{aligned} \text{rect}(x) &= 1 , \quad |x| \leq \tfrac{1}{2} \\ &= 0 , \quad |x| > \tfrac{1}{2} \end{aligned} \tag{1.9}$$

and I_+ is the set of positive integers.

If h is of a simple (cf. Table 1.1) but unknown form, it is possible to determine the particular simple form of h from the zeros of $|F(u,v)|$ (Gennery 1973). Taking the FT of (1.6), neglecting noise, and using the convolution theorem for Fourier transforms gives

$$F = PH. \tag{1.10}$$

It follows that the set of zeros of F is the union of the set of zeros of P and the set of zeros of H. If h has a simple form, like any of those in Table 1.1, the zeros of H form a simple pattern (e.g. they lie along equispaced parallel straight lines for uniform motion blur, or on rings for out-of-focus blur). In general the zeros of P do not form a simple pattern. Therefore, the pattern of the zeros of H is clearly visible among the zeros of F, if $|F|$ is displayed, so that the form of h may be deduced. For simple blurs, the extent of h may be calculated from the spacing of the zeros (Gennery 1973).

1.3.2 Edge Analysis

Alternatively, it is possible to deduce h by examining blurred edges in f, when it is reasonable to assume that the edges would have been sharp in the absence of degradations (NASA/ERC Seminar 1968, Huang *et al* 1971, Tescher and Andrews 1972). To show how this is done let an

edge be modelled as a unit step function step(x) defined by

$$\begin{aligned} \text{step}(x) &= 0, \quad x < 0, \\ &= 1, \quad x \geqslant 0, \end{aligned} \tag{1.11}$$

and consider this edge as the original image. Neglecting noise, f is given by

$$f(x,y) = \text{step}(x) \otimes h(x,y) \tag{1.12}$$

$$= \int_{-\infty}^{\infty}\int \text{step}(x-\alpha)h(\alpha,\beta)d\alpha d\beta. \tag{1.13}$$

Now

$$\frac{\partial f(x,0)}{\partial x} = \int_{-\infty}^{\infty} h(x,\beta)d\beta \tag{1.14}$$

which is the projection of h parallel to the y-axis (Smith et al 1973). Edges in different directions give different projections of h. An estimate of h may be reconstructed from a finite number of these projections (Smith et al 1973). If h is circularly symmetric it may be recovered from a single projection. If h is one-dimensional (1-D) then (1.13) reduces to

$$\frac{\partial f(x,0)}{\partial x} = h(x), \tag{1.15}$$

so that h may be found directly.

1.3.3 Image Segmentation

An interesting method of estimating H, and hence h, from a large degraded image is discussed by Huang et al

(1971). It is based on work done by Prof. T.G. Stockham on the restoration of old Caruso recordings. The results of this investigation were later presented in detail (Stockham et al 1975).

The final degraded image is divided into M segments of equal size, and the intensity distribution in the jth segment is denoted by $f^j = f^j(x,y)$, where $j = 1,\ldots,M$. The corresponding intensity distribution of the original image in the jth segment is denoted by $p^j = p^j(x,y)$. The effect of convolving p^j with h is not confined to the jth segment because the degraded image is spread into adjacent segments. This effect may be reduced by multiplying each segment by a window which falls off to zero at the edge of each segment (Stockham et al 1975). It is then approximately true that

$$f^j = p^j \otimes h , \quad j = 1,\ldots, M. \tag{1.16}$$

so that (1.10) gives

$$F^j = p^j H , \quad j = 1,\ldots, M. \tag{1.17}$$

Taking the product of (1.17) over j and solving for $|H|$ yields

$$|H| = \left| \left[\prod_{j=1}^{M} F^j \right]^{\frac{1}{M}} \right| \Bigg/ \left| \left[\prod_{j=1}^{M} p^j \right]^{\frac{1}{M}} \right| . \tag{1.18}$$

An a priori guess is made at the form of $\left| \left[\prod_{j=1}^{M} p^j \right]^{\frac{1}{M}} \right|$ so that (1.18) gives an estimate of $|H|$. For simple blurs (as in Table 1.1) an estimate of h may then be obtained.

A refinement of this technique is suggested in chapter 8.

1.3.4 Calibration Techniques

One method of determining h which comes under the heading of calibration techniques is proposed by Ekstrom (1973a). He considers the case when the degraded image f of a known original image p is available. Neglecting noise and using a discrete model of (1.6) he obtains

$$\underset{\sim}{f} = \underset{\sim}{P}\,\underset{\sim}{h} \tag{1.19}$$

where (in matrix notation) $\underset{\sim}{f} = (f_j)$, $\underset{\sim}{h} = (h_k)$ and $\underset{\sim}{P} = (p_{j,k})$ for $j = 1,\ldots,N_1$ and $k = 1,\ldots,N_2$ and where $N_1 > N_2$. In matrix notation, the convolution in (1.6) corresponds to multiplication of $\underset{\sim}{h}$ by a matrix $\underset{\sim}{P}$ of dimension $N_1 \times N_2$. (A 1-D explanation is given here for simplicity). The method of least squares is then used to obtain an estimate of $\underset{\sim}{h}$ given by

$$\underset{\sim}{h} = (\underset{\sim}{P}^t\underset{\sim}{P})^{-1}\underset{\sim}{P}^t\underset{\sim}{f} \tag{1.20}$$

where the superscript t denotes the matrix transpose. Methods such as this are often used when calibration is possible. The edge detection methods described in 1.3.2 are the most commonly used when h cannot be determined through calibration. The advantage of the method described in 1.3.3 is that it needs no edges and is always applicable if the extent of the psf is considerably less than that of the blurred image.

1.4 DIGITAL IMAGE RESTORATION

Image restoration methods for obtaining p from $\underline{f}$, as given by (1.6), are in general either optical or digital. Optical image restoration is performed by a combination of holography and standard photographic and optical techniques. Alternatively, the image can be digitised by scanning it with a computer-controlled microdensitometer, so that it can be stored as discrete "pixels" (picture elements) in computer memory - the restoration is then carried out in a digital computer. Digital and optical techniques are reviewed separately in 1.4 and 1.5 respectively, and are compared in 1.6.

Digital image restoration has recently been reviewed by Huang et al (1971), Sondhi (1972), Andrews (1974) and Hunt (1975). It has also featured in a number of special issues of journals (e.g. Proc. IEEE July 1972, Computer May 1974, and Optical Eng. May-June 1974). Two introductory books have been published on the subject (Rosenfeld 1969, and Andrews 1970).

Various digital image processing laboratories are in operation and have periodically issued reports describing their facilities and the tasks they carry out (Patterson et al 1974, Hunt et al 1974, Billingsley 1973).

Digital restoration techniques are broadly divisible into four categories: recursive, nonrecursive, positive and superresolution, and miscellaneous. These are dealt with in turn.

It is convenient here to introduce another division,

due to Andrews (1974), of restoration techniques into two categories: continuous/continuous methods and discrete/discrete methods. In the continuous/continuous methods, both f and p are considered to be continuous. It is then assumed that f is bandlimited, so that the continuous f may be obtained from the samples of f taken at the Nyquist rate (Goodman 1968). It is said that f is band-limited to a frequency N_f if

$$|F(u,v)| = 0 \quad \text{for } \sqrt{u^2+v^2} > N_f, \tag{1.21}$$

in 2-D, or

$$|F(u)| = 0 \quad \text{for } |u| > N_f, \tag{1.22}$$

in 1-D (i.e. for f = f(x)). The Nyquist sampling rate in a given direction is equal to the reciprocal of the actual (or effective) extent of F in that direction. In the discrete/discrete methods it is assumed that both f and p are sampled and discrete.

1.4.1 Recursive Techniques

A recursive digital filter is one which gives a sequence of output samples as a weighted sum of input samples and previously calculated output samples. Such a filter may be used to give samples of $\hat{p}$ from samples of f and previously calculated samples of $\hat{p}$. In 1-D, if $\hat{p}_k$ and f_k (where $k \epsilon I_+$) are samples of $\hat{p} = \hat{p}(x)$ and f = f(x) respectively, then in general the $\hat{p}_k$ are given by

$$\hat{p}_k = \sum_{l=1}^{M_1} c_l \hat{p}_{k-l} + \sum_{j=-M_2}^{M_2} d_j f_{k-j}, \tag{1.23}$$

where M_1, M_2, the c_l and the d_j are constants. Such filters are widely used in 1-D signal processing (IEEE Selected Reprints on Digital Signal Processing 1972). A number of approaches have been taken to the problem of extending these filters to 2-D (e.g. Shanks 1970). Some attempts have been made to apply these 2-D recursive filters to image restoration (Zimmermann and Gupta 1973, Hall 1972). However, their usefulness has been limited for a number of reasons. Firstly, because of the lack of a 2-D factorisation theorem (Huang et al 1971), many results for 1-D recursive filters do not hold in 2-D. This makes it difficult to approximate a desired frequency response with a stable 2-D filter. Secondly, to be effectively implemented, a 2-D recursive filter needs to be approximated by a finite sum of separable filters (Huang et al 1971). A third problem is the choice of boundary values needed to start the recursion. Because the number of sample points in an image is often not large (e.g. 128 x 128 pixels), the effect of errors in the chosen boundary values may spread throughout the whole image. For these reasons, recursive filters are not much used in image restoration at present.

1.4.2 Nonrecursive Techniques

The use of nonrecursive filters for image restoration is widespread. The outputs of such filters depend only on the input samples. Unlike with nonrecursive filters, there is no feedback of previously calculated outputs as in (1.23). Corresponding to (1.23) the general expression for the output of a nonrecursive filter in 1-D is

$$\hat{p}_k = \sum_{j=-M_2}^{M_2} d_j \, f_{k-j}, \qquad k \in I_+, \tag{1.24}$$

where M_2 and the d_j are constants. The main advantages of these filters are that there is no stability problem, they are easily implemented, noise analysis is relatively simple, and the extent of the resultant psf (i.e. the result of convolving the psf with the filter) in the restored image can be conveniently chosen. Nonrecursive filters are applied either by convolution in image space $\mathcal{I}$ or by multiplication in Fourier space $\mathcal{F}$.

1.4.2.1 Fourier Methods

Nonrecursive filtering is carried out in $\mathcal{F}$ by means of the Fourier or the Hadamard transforms. Such transformations have been made practical by the development of efficient algorithms for carrying them out on the digital computer. The Fast Fourier Transform algorithm or FFT was put forward by Cooley and Tukey in 1965. Later a fast Hadamard transform also became available (Andrews 1970, Robinson 1972). Only the FFT is considered here because similar results hold for both transforms, and the FFT is the one more widely in use. Other fast transforms such as the Fermat number transform (Agarwal and Burrus 1974) also exist.

Unlike the FT in (1.8) which acts on a continuous function f, the FFT acts on a rectangular array $\underset{\sim}{f}_s = \{f_{k,l};\ k,l=1,..,N\}$ of samples of f. For convenience the array is considered to be square. The FFT produces a sampled FT distribution $\underset{\sim}{F}_s = \{F_{m,n};\ m,n=1,\ldots,N\}$,

where

$$F_{m,n} = \frac{1}{N^2} \sum_{k=1}^{N} \sum_{l=1}^{N} f_{k,l} e^{2\pi i(km+ln)/N^2}. \qquad (1.25)$$

Various Fourier transform filtering techniques which have become known as "inverse filtering" are now discussed. In line with most work on the subject (Huang et al 1971, Andrews 1974) the filtering equations are presented in continuous form, although their digital implementation by means of the FFT is necessarily discrete.

Taking the FT of (1.6) gives

$$F = P H + N. \qquad (1.26)$$

In the procedure known as inverse filtering, $\hat{p}$ is taken to be the inverse FT of $F\overline{H}$, where $\overline{H} = \overline{H}(u,v)$ is called the "inverse filter" for the following reason. When $n = 0$ it follows from (1.26) that $\hat{p} = p$ if

$$\overline{H} = 1/H, \qquad (1.27)$$

which is known as the direct inverse filter.

Two shortcomings of the direct inverse filter (1.27) are as follows. Firstly, H may have zeros at spatial frequencies within the range of interest (e.g. uniform motion blur and out-of-focus blur), in which case 1/H does not exist. Secondly, division by H tends to unduly amplify any noise that is present at or near zeros of H. This noise amplification tends to increase with frequency because $|H|$ generally decreases with increasing frequency. These problems are discussed by Sondhi (1972), Huang et al (1971),

Andrews 1974) and Hunt (1975). A number of authors (Harris 1966, McGlamery 1967, Mueller and Reynolds 1967) describe restoration techniques using filters which are minor modifications of (1.27). For example, Nathan (1971) uses (1.27) with the proviso that $\overline{H}$ is not allowed to exceed a set constant.

When $n \neq 0$, which is always true in practice, perhaps the most suitable form for $\overline{H}$ is the Wiener filter (Helstrom 1967, Slepian 1967a, Horner 1969)

$$\overline{H} = \frac{H^*}{H H^* + \Phi^2}, \tag{1.28}$$

where the asterisk denotes the complex conjugate and $\Phi = \Phi(u,v)$ is an a priori estimate of the ratio of $|N|$ to $|P|$. The Wiener filter is optimum in the sense that it results in a $\hat{p}$ for which

$$\int_{-\infty}^{\infty}\int |\hat{p}(x,y) - p(x,y)|^2 \, dxdy \tag{1.29}$$

is minimised (Sondhi 1972). A number of modifications to the Wiener filter have been proposed (Hunt 1975, Cole 1973). For example, Cole (1973) suggests a geometric mean filter which can be classed between the inverse filter and the Wiener filter:

$$\overline{H} = \left(\frac{1}{H}\right)^{s'} \left(\frac{H^*}{|H|^2 + \Phi^2}\right)^{1-s'} \tag{1.30}$$

where $0 \leq s' \leq 1$. Gennery (1973) simplifies (1.28) by assuming Φ to be a constant equal to the standard deviation

of the noise which he then estimates.

Constrained least squares methods are discussed in 1.4.4.1 for the SV case. The corresponding filter for the SI case is (Hunt 1975)

$$\overline{H} = \frac{H^*}{H H^* + \lambda \mathscr{C}\mathscr{C}^*} \tag{1.31}$$

where λ is a parameter determined by iteration and $\mathscr{C}(u,v)\mathscr{C}^*(u,v)$ corresponds to $\underset{\sim}{C}$ in (1.50) (refer to 1.4.4.1).

The MSE criterion in (1.28) for the optimum $\hat{p}$ is not necessarily the best. Optimality criteria are discussed by Sondhi (1972) and Hunt (1975). The restoration procedure yielding the result $\hat{p}$ which appears the best, depends on f, the detail of interest in p, and also on the interpretative powers of the human observer. Thus, human criteria for judging the best restoration are essentially subjective. However, to be practically implemented, criteria need to be objective. The advantages of the MSE criterion are that it is objective and easily computable. The problem of presenting an image in the most suitable form for a human observer is largely an enhancement problem, for which techniques such as histogram equalisation (Andrews 1974) can be used. Even so, it is desirable that the optimality criteria take into account some of the eye's properties. For example, the eye places more emphasis on mid-frequency information than it does on either high or low frequency information (Hunt 1975). The following filter has been suggested by Hunt (1975) to account for the properties of the eye:

$$\bar{H} = \frac{H^*}{H H^* + \lambda/(|H_E|^2)} \qquad (1.32)$$

where $|H_e| = |H_e(u,v)|$ is the modulation transfer function of the eye. Nevertheless, the MSE criterion has been widely adopted and is shown to produce good restorations by a number of authors (Gennery 1973, McGlamery 1967, Harris 1968).

1.4.2.2 Direct Methods

Nonrecursive filtering can be accomplished directly by convolution with a finite filter array in $\mathcal{I}$ as in (1.24). One advantage this offers is that it is faster than convolution using the FFT, if the array is small enough (Hall 1972). The main disadvantage is that direct methods lack the versatility of Fourier methods for dealing with noise. Implementation of direct nonrecursive filters can be either digital, or optical, using incoherent light to superimpose displaced negatives (Frieden 1974), or electro-optical by scanning with a specially designed aperture and writing the result directly on to film (Smith 1966, Honda et al 1974, 1975).

There are a number of signal processing techniques in use for the design of finite impulse response (FIR) filters in 1-D (IEEE Selected Reports on Digital Signal Processing 1972). The aim of these techniques, however, is usually to approximate a desired frequency response (e.g. a band pass filter (Rabiner 1971)). In image restoration the aim of the design is to make the resultant psf as close as possible in some sense to a delta function. This problem is considered

in 1-D by Smith (1966) when designing a continuous filter rather than a filter array. He minimises the radius of gyration of the resultant psf. However, the solution for a filter array is more useful and this is considered in both one and two dimensions by a number of authors (Saleh 1974, Frieden 1974, Stuller 1972, Riemer and McGillem 1973, Nathan 1971, Arguello et al 1972, and Honda et al 1974, 1975). The various techniques they propose for deriving the filter array and dealing with noise are considered in detail in chapter 6. It suffices to say here that nonrecursive filter arrays provide a viable alternative to Fourier filtering methods.

1.4.3 Positive Restoration and Superresolution

Positive restoration methods make use of (1.7) in addition to (1.6) so that the solution is generally more difficult than for the methods previously mentioned. Superresolution aims at recovering the spatial frequency content of an image that has been lost through bandlimiting. Positive restoration is considered here with superresolution because positive restoration methods can be used to obtain superresolution. Positive restoration and superresolution are discussed by Andrews (1972, 1974), Gerchberg (1974) and Huang et al (1971).

1.4.3.1 Positive Restoration

Positive restoration methods have only recently become available. They have been used with some success in 1-D applications, but their use is not yet practical in 2-D.

Biraud (1969) develops a positive restoration

technique to achieve superresolution in 1-D. He assumes that $P = P(u)$ is given up to some cutoff frequency u_c, and wishes to extend P to a new cutoff frequency $u_c' > u_c$. As $p = p(x)$ is positive he defines a new function $g = g(x)$ by

$$p(x) = (g(x))^2 \geq 0 , \tag{1.33}$$

so that

$$P(u) = G(u) \otimes G(u). \tag{1.34}$$

He then develops an iterative technique to find an estimate $\hat{G} = \hat{G}(u)$ of G which minimises

$$\int_0^{u_c} |P(u) - \hat{P}(u)|^2 du, \tag{1.35}$$

where

$$\hat{P}(u) = \hat{G}(u) \otimes \hat{G}(u). \tag{1.36}$$

$\hat{P} = \hat{P}(u)$ given by (1.36) is bandlimited to u_c' as required.

Other authors (Philip 1963, 1973, Frieden 1972, Frieden and Burke 1972, Lahart 1974) use a maximum likelihood or maximum entropy restoration which finds the most likely original image given a degraded image formed from a statistical distribution. Alternatively, the image may be assumed to be a 2-D probability density function, so that Bayes theorem can be used for the restoration (Richardson 1972).

Frieden (1973) suggests an interesting 1-D technique in which $\hat{p}$ is built up by the cumulative placement of fixed

intensity increments or "grains" at samples of $\hat{p}$. A corresponding estimate $\hat{f} = \hat{f}(x)$ of the degraded image is cumulatively built up and at any stage is given by

$$\hat{f} = \hat{p} \otimes h. \tag{1.37}$$

A decision rule is required for the placement of grains. A grain of fixed intensity d_o is placed at a sample x_n of $\hat{p}$ if x_n permits the inequality

$$\hat{f}(x_m) + d_o h(x_m - x_n) \leqslant r_m \hat{f}(x_m) \tag{1.38}$$

to be obeyed for all samples x_m, for a minimum value of r_m. This method gave encouraging results in a computer simulation, but is not yet practical for use in 2-D.

Another technique which is classified as positive by Andrews (1974) is the modified Van Cittert method used by Jansson (1970) and Jansson et al (1970). Neglecting noise, the 1-D form of (1.6) may be written as a matrix equation

$$\underset{\sim}{f} = \underset{\sim}{H}\,\underset{\sim}{p} \tag{1.39}$$

where in matrix notation $\underset{\sim}{f} = (f_j)$ and $\underset{\sim}{p} = (p_k)$ are column rectors of length N, and $\underset{\sim}{H} = (h_{j,k})$ is a square matrix. Attempts to solve (1.39) directly by inversion of $\underset{\sim}{H}$ to yield

$$\underset{\sim}{p} = \underset{\sim}{H}^{-1}\underset{\sim}{f} \tag{1.40}$$

lead to unacceptable solutions because of the instability of $\underset{\sim}{H}$ to inversion. An iterative technique is therefore applied to give estimates $\hat{\underset{\sim}{p}}^k = (\hat{p}_j^k)$, where k = 1, 2.., of $\underset{\sim}{p}$.

The p_j^{k+1} are obtained from

$$\hat{p}_j^{k+1} = \hat{p}_j^k + \frac{\kappa}{h_{j,j}} \{f_j - \sum_{l=1}^{N} h_{j,l} \hat{p}_l^k\} \qquad (1.41)$$

where κ is a relaxation parameter and $\hat{\underset{\sim}{p}}^1 = \underset{\sim}{f}$. This approach has given good results when applied to physical spectra in 1-D. However, it would be unwieldy in 2-D because of the size of the matrices that would be involved.

1.4.3.2 Superresolution

Analytic continuation is the principal method of obtaining superresolution. The basis of the approach is that the FT of an image of finite extent is analytic, so that the whole FT can be obtained, in the absence of noise, from any portion of it. Prolate spheroidal wave functions (PSWF's) are ideally suited to analytic continuation (Slepian and Pollak 1961). They are used for superresolution by a number of authors (Frieden 1967, Barnes 1966, Rino 1969, Huang et al 1971). An earlier technique not using PSWF's was discussed by Harris (1964). The main limitation of these methods is that they are very unstable in the presence of noise and only a very small extension of the FT seems to be practical (Rushforth and Harris 1968). Once again experimental results have been restricted to 1-D.

It seems clear that the iterative approach of Gerchberg (1974) is the one most likely to be practical for superresolution in 2-D. He views (in 1-D) the bandlimited spectrum $P_B(u)$ as the sum of the true spectrum $P(u)$ and an error spectrum $R(u)$, both of which are not bandlimited.

It is assumed that the extent of p(x) is less than or equal to A. The (j+1)th estimate $P_{j+1}(u)$ of P(u) is obtained from $P_j(u)$ as follows. Take the inverse FT of $P_j(u)$ to obtain $p_j(x)$. Obtain $p_{j+1}(x)$ from

$$p_{j+1}(x) = p_j(x)\ \mathrm{rect}(x/A). \tag{1.42}$$

Then $P_{j+1}(u)$ is given by

$$\left.\begin{aligned} P_{j+1}(u) &= P_B(u), \qquad u \leqslant u_c, \\ &= \text{FT of } p_{j+1}(x), \qquad u > u_c. \end{aligned}\right\} \tag{1.43}$$

The resultant error after each iteration $R_{j+1}(u)$ is given by

$$P_{j+1}(u) = P(u) + R_{j+1}(u). \tag{1.44}$$

Gerchberg shows, using Parseval's theorem, that the 'energy' of the error spectrum is monotonically decreasing as j increases. The technique is surprisingly stable to noise and the results are limited more by the accuracy of the estimate of the extent of the image, than by the noise level present. De Santis et al (1975) have reformulated Gerchberg's algorithm using PSWF's.

1.4.4 Miscellaneous Methods

1.4.4.1 Matrix Methods

In 1-D (1.4) may be written using (1.1) as

$$f(x) = \int_{-\infty}^{\infty} h(x,\alpha)p(\alpha)d\alpha + n(x), \tag{1.45}$$

which is in the form of a Fredholm integral equation of the first kind. The discrete/discrete version of (1.45) is

$$\underset{\sim}{f} = \underset{\sim}{H}\,\underset{\sim}{p} + \underset{\sim}{n} , \qquad (1.46)$$

where (in matrix notation) $\underset{\sim}{f} = (f_j)$, $\underset{\sim}{p} = (p_k)$, $\underset{\sim}{n} = (n_k)$, $\underset{\sim}{H} = (h_{j,k})$ for $j = 1,..,N$ and $k = 1,..,M$. A number of numerical analysis techniques have been proposed for solving this equation for $\hat{\underset{\sim}{p}} = (\hat{p}_j)$, $j = 1,..,N$. As written, (1.46) describes an SV system. When h is SI, the Fourier inverse filtering techniques of 1.4.2.1, using the FFT, become applicable with a consequent increase in processing speed.

In continuous form (1.45) has no stable solution because a high frequency oscillation in p can produce a negligible effect on f (Phillips 1962, Sondhi 1972). If f and p are sampled at rates such that N = M, if n = 0, and if the determinant of $\underset{\sim}{H}$ is nonzero, then in principle (1.46) is solvable for $\hat{\underset{\sim}{p}}$. However, in practice, because of the size of N, inversion of $\underset{\sim}{H}$ is slow, and $\underset{\sim}{H}$ is almost singular (Hunt 1972, Ekstrom 1973b,c). This makes the solution very unstable in the presence of noise. If (1.1) is expressed in matrix form in 2-D, with the 2-D matrices being rewritten as 1-D ones (MacAdam 1970, Hunt 1973), the problems become even worse (Andrews 1974, Huang et al 1971).

As well as the technique of Jansson et al (1970) characterised by (1.41), pseudoinverse techniques (Adler et al 1974) have been suggested for obtaining $\hat{\underset{\sim}{p}}$ when $\underset{\sim}{H}$ is almost singular. Another approach to the matrix inversion problem (Andrews 1974) is given by the power series method

where

$$\underset{\sim}{\hat{p}}^{k} = \sum_{l=0}^{k} (\underset{\sim}{I}-\underset{\sim}{H})^{l} \underset{\sim}{f} \tag{1.47}$$

or

$$\underset{\sim}{\hat{p}}^{k} = \underset{\sim}{f} + (\underset{\sim}{I}-\underset{\sim}{H})\underset{\sim}{\hat{p}}^{k-1}, \tag{1.48}$$

where $\underset{\sim}{I}$ is the identity matrix. Such iterations, although slow, are capable of giving, in 1-D, a $\underset{\sim}{\hat{p}}$ which is a considerable improvement on $\underset{\sim}{f}$.

It is possible to solve (1.46) by placing constraints on it. Phillips (1962) proposes that the $\underset{\sim}{\hat{p}}$ be found which is the smoothest solution in the sense that

$$\int_{-\infty}^{\infty} \left(\frac{\partial^2 \hat{p}(x)}{\partial x^2}\right)^2 dx \tag{1.49}$$

is minimised. Twomey (1963) gives a more generalised solution which does not require $\underset{\sim}{H}$ to be square. He puts a constraint on the noise and the smoothness, obtaining

$$\underset{\sim}{\hat{p}} = (\underset{\sim}{H}^{t}\underset{\sim}{H} + \mu \underset{\sim}{C})^{-1} \underset{\sim}{H}^{t}\underset{\sim}{f}, \tag{1.50}$$

where $C_p = \underset{\sim}{p}^{t}\underset{\sim}{C}\underset{\sim}{p}$ is minimised, where $\underset{\sim}{C}$ is some square matrix that makes C_p a reasonable criterion of smoothness. μ is chosen by trial and error to give the best solution. Setting $\mu = 0$ in (1.50) results in

$$\underset{\sim}{\hat{p}} = (\underset{\sim}{H}^{t}\underset{\sim}{H})^{-1} \underset{\sim}{H}^{t}\underset{\sim}{f}, \tag{1.51}$$

which is a direct inverse filter generalised for the SV case. Corresponding generalisations of the Wiener and other filters may also be found (Andrews 1974, Sondhi 1972). All of the matrix methods discussed in this section suffer from

two major disadvantages, namely lack of speed and impracticability in 2-D.

MacAdam (1970) shows how any 2-D convolution can be rewritten as a 1-D convolution. He then develops an algorithm for doing deconvolution under the constraint that the values of $\hat{p}$, f, and h lie within prescribed ranges. He suggests that the constraints should be modified after each attempt at a restoration. The danger here is that the programmer may end up partly drawing his own picture. The main disadvantage of the method is its slowness, if f has a large number of samples.

1.4.4.2 Other Methods

A technique which has found some application in 1-D image restoration (McKinnon et al 1975, Bates et al 1975) is based on the following result. A function of finite extent is characterised uniquely by the zeros of the complex FT (Bates 1969). The complex FT is defined as in (1.8) except that u and v are complex. The basic result is that the set of complex zeros of b in (1.1) is the union of the sets of complex zeros of p and h. Thus, deconvolution may be performed by subtracting the set of complex zeros of h from that of b, and then obtaining the corresponding $\hat{p}$ in image space.

A number of authors consider the case when $h(x,y,\alpha,\beta)$ is not deterministic, but a random function of position (α,β) with known statistics. The results which they derive involve minor modifications of standard Fourier inverse filtering theory (Slepian 1967a, Sondhi 1972).

Another possibility known as multiframe processing is that a number of images of a single object may be available, each of which has been degraded by a different psf. The various methods proposed for dealing with this case are discussed by Huang et al (1971). However, the main interest in this thesis is with a single image degraded by a deterministic psf.

The restoration techniques discussed so far are linear. However, nonlinear techniques for restoring linearly degraded images are also possible. Frieden (1968) derives a filter to minimise

$$\int_{-\infty}^{\infty}\int |\hat{p}(x,y) - p(x,y)|^{K} dx\, dy, \qquad (1.52)$$

instead of (1.29), where K is a positive integer. A thorough treatment of nonlinear filtering techniques is given by Oppenheim et al (1968). Their use is demonstrated to be practicable by Stockham et al (1975). However, non-linear techniques, though potentially more powerful than linear ones, have not yet found wide application because the slight processing advantages they offer are outweighed by their increased complexity and decreased speed.

In (1.4) the noise present in f is assumed to be additive. Although this is the simplest and most commonly used model, it is not necessarily the most accurate one. Some types of noise are better modelled as being multiplicative (Huang 1966, Hunt 1975). Both Stockham (1972) and Hunt (1975) consider ways of dealing with multiplicative noise, and Anderson et al (1971) consider

the effects of various types of noise in a system. However, at present, image restoration quality seems to be limited more by the image restoration technique used than by the inaccurate model of additive noise.

In each restoration technique which has been discussed, except for direct nonrecursive filtering and the slow matrix methods in 1.4.4.1, it is effectively assumed that all of f given by (1.4) is available. However, any real-world image must be of finite size because it exists within a finite recording frame. As is explained in detail in chapter 2, it makes sense to consider the special class $\mathcal{S}$ of degraded images which fit completely inside their recording frames; and the general class $\mathcal{G}$ of degraded images which are truncated by their recording frames. In the case of class $\mathcal{G}$ images, the restoration techniques which have been discussed (except for direct nonrecursive filtering and the slow matrix methods in 1.4.4.1) are unsoundly based and may give unacceptable results.

This serious problem is mentioned but not solved by Huang et al (1971), Ekstrom (1973c), Hoppe (1970), Lewis et al (1975), and Campbell et al (1974). It is discussed in detail in chapter 2. In almost all published restorations (Harris 1966, McGlamery 1967, MacAdam 1970, Biraud 1969, Robbins 1970, Sondhi 1972, Hunt 1973, 1975, Inuiya et al 1973) the degraded image is of class $\mathcal{S}$ (where the background intensity level is assumed to be zero). In the particular case of direct nonrecursive filtering, or an equivalent FIR filter implemented using the FFT (Arquello et al 1972), the

truncation of f because it is of class $\mathcal{G}$ only has an effect on a band at the edge of $\hat{p}$, whose width equals the width of the filter array. This is noted by Ekstrom (1973c).

1.5 OPTICAL IMAGE RESTORATION

Vander Lugt (1968), Stroke (1972), Huang et al (1971) and Campbell et al (1974) are among the many who have reviewed optical image restoration techniques, which are usually effected with a coherent optical system similar to that shown in Fig. 1.2.

The FT property (Goodman 1968) of a lens means that if a film transparency with amplitude transmittivity f(x,y) is placed in the front focal plane (a) of lens (b), and is illuminated with a collimated uniform beam of coherent monochromatic light, then the amplitude at the back focal plane (c) of lens (b) is

$$F\left(\frac{2\pi\xi^1}{\lambda f_1}, \frac{2\pi\eta^1}{\lambda f_1}\right) = F(u,v) \qquad (1.53)$$

where (ξ^1,η^1) are the spatial coordinates of the back focal plane, λ is the wavelength of the light, and f_1 is the focal length of each lens. This means that inverse filtering may be performed optically simply by placing a film transparency, whose amplitude transmittivity is equal to the desired filter $\overline{H}$, in the back focal plane (c) of lens (b) (Goodman 1968, Stroke 1972). A second lens (d) is then used to perform the inverse FT to give $\hat{p}$ in plane (e).

Most optical filters in practical use are gross

approximations of those described in 1.4.2.1, because the dynamic range of any transmissive medium is limited. It is not possible to make an optical filter which has a transmittivity greater than unity. When a filter is, as it usually is, complex it must be produced by a holographic method or by using special phase retarding plates (Maréchal et al 1953). Holographic methods have proven simpler and more accurate (Vander Lugt 1968, Stroke 1972, Lee 1970, Campbell et al 1974). Holographic filters have density variation only and may be formed optically (Stroke 1972), or by means of a digital computer (Gough and Bates 1972). Optical filtering techniques have been used to obtain successful restorations of class $\mathscr{S}$ images under laboratory conditions (Stroke and Halioua 1973a, 1973b, Lohmann and Werlich 1967, Lui and Gallagher 1974). Campbell et al (1974) have attempted to restore a class $\mathscr{G}$ image optically. They find that artefacts are present throughout $\hat{p}$. These are caused by the convolution of the filter impulse response, $\bar{h}$, with the sharp edges of the truncated degraded image.

Optical image restoration may also be carried out using incoherent light (Frieden 1974, Honda et al 1974, 1975). For example, Frieden (1974) suggests that an image may be restored by superimposing displaced negatives. These methods are dealt with in 1.4.2.2 because they are mathematically equivalent to direct nonrecursive filtering techniques.

1.6 COMPARISON OF OPTICAL AND DIGITAL METHODS

Digital and optical image restoration techniques are compared by Huang et al (1971) and Stroke (1972).

The main advantage of digital restoration is its versatility. For example, it is simple to vary the filter interactively, whereas even small changes in an optical filter may require that a new filter be constructed (this may take days). Also, digital restoration may be used to do a number of nonlinear operations which have no optical counterpart, although Lee (1974) discusses nonlinear operations which can be achieved optically.

The great advantage of optical restoration is the speed with which an image may be processed, and the large amount of information that can be stored simultaneously. Because optical processing is parallel, the speed with which an image may be restored is limited only by the time taken to position transparencies and film in an optical bench. The enormous information storage capacity of photographic film means that a large number of resolvable points are processed each time. On the other hand, digital techniques are typically applied to a sampled image consisting of as few as 128 x 128 pixels. The time taken for a digital restoration using, for example, a Burroughs B6718 computer is commonly of the order of two minutes. An advantage of digital restoration is its accuracy and reliability. Inaccuracies in the restoration are due to quantisation errors in sampling and processing, and to errors in estimating the characteristics of the degradation.

The latter error necessarily effects both optical and digital techniques similarly. Errors due to sampling are generally of the order of 1%, but may be made as small as 0.01% with sophisticated scanning equipment. In general, processing errors are negligible. There are a number of error sources in optical restoration other than those due to estimating the characteristics of the degradation (Huang et al 1971). These are - imperfect optical components, film grain noise, nonlinearities, and more importantly: errors in constructing the filter, and the presence of speckle noise (characteristic of coherent systems) on the restored image. These are serious practical limitations on the use of optical image restoration methods.

Both optical and digital restoration methods are expensive. An optical image processing system is expensive because of the cost of a high quality coherent system. A digital image processing system is expensive because of the cost of a computer, but this may be hired. Actual operating costs are probably higher for digital methods because of the cost of computer time.

Overall it is my opinion that the digital advantages of versatility and accuracy outweigh the optical advantages of speed and capacity. This is especially true when class $\mathcal{G}$ images are being restored. Optical image restoration seems preferable when a large number of class $\mathcal{A}$ images are to be restored with a single filter.

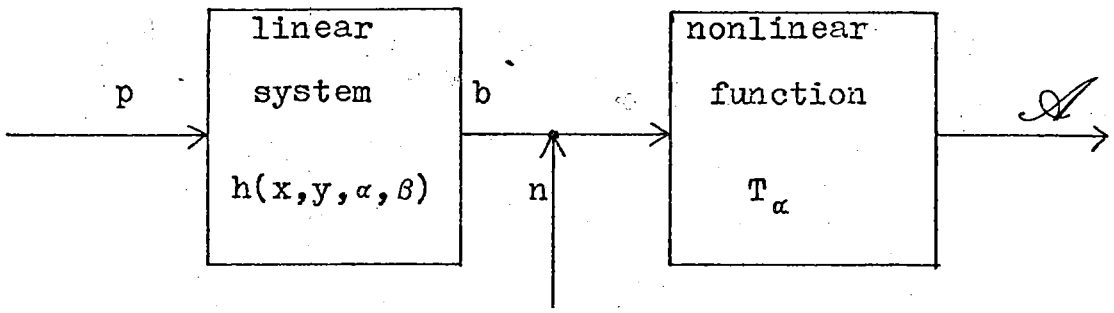


Fig.1.1. Model for a degrading system.

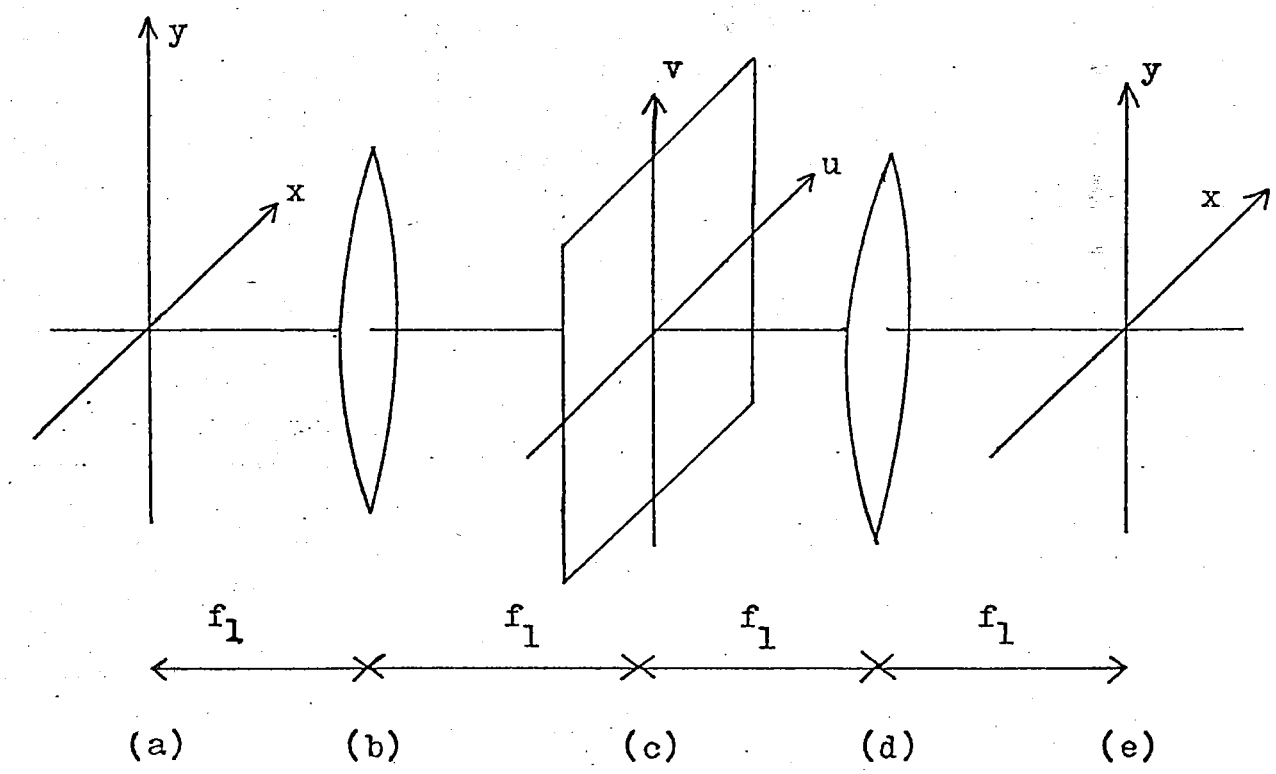
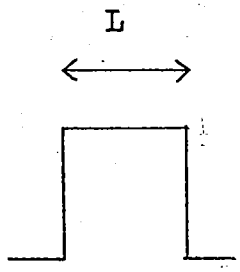
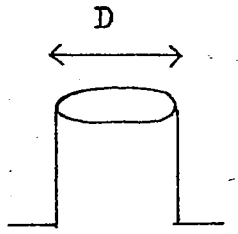
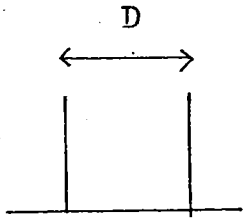


Fig.1.2. A coherent system for optical filtering.

Table 1.1. Some common forms of blurring.

type of blur	form of h	h	H	positive zeros of H
uniform motion		$\frac{1}{L}\,\mathrm{rect}(x/L)$	$H(u) = \mathrm{sinc}(Lu)$	$u = m/L,$ $m \in I_+$
out-of-focus		$h(r;\theta) = 0, r > \frac{D}{2}$ $= \frac{4}{\pi D^2}, r \leq \frac{D}{2}$	$H(\rho;\phi) = \frac{2J_1(\pi\rho D)}{\pi\rho D}$	$\rho = \frac{j_{1,m}}{\pi D},$ $m \in I_+,$ $J_1(j_{1,m}) = 0$
overlay		$h(x) = (\delta(x-D/2) + \delta(x+D/2))/2$	$H(u) = \cos(\pi u D)$	$u = \frac{2m-1}{2D},$ $m \in I_+$

CHAPTER TWO

COMPREHENSIVE NOTATION AND DISCUSSION OF SOME IMAGE RESTORATION PROBLEMS

2.1 A COMPREHENSIVE NOTATION FOR IMAGE RESTORATION

In this section, notation is introduced in addition to that already introduced in chapter 1. In particular, various frames in image space are defined, as well as the concepts of class $\mathscr{S}$ and class $\mathscr{G}$ images. The insights provided by this new notation are basic to an understanding of the results presented in chapters 3, 4 and 5. Some of the definitions already given in chapter 1 are repeated here for clarity.

As well as image space $\mathscr{I}$, in which images are represented by non-negative real functions, Fourier space $\mathscr{F}$ is required. As indicated in Fig. 2.1, γ denotes an arbitrary point, with Cartesian coordinates x and y, in $\mathscr{I}$; and ω denotes an arbitrary point, with Cartesian coordinates u and v in $\mathscr{F}$. The final degraded image appearing within the recording frame Γ, which is a simply connected domain of $\mathscr{I}$, is denoted by $f = f(x,y)$. The original image, which it is hoped to recover from f, is denoted by $p = p(x,y)$.

The degradation of the image is considered to take place in two steps. There is the preliminary degradation, which is the distortion (denoted here by $\mathscr{D} = \mathscr{D}\{p\}$) introduced by the physical process responsible for degrading

the original image. Included in $\mathscr{D}$ is any remaining nonlinear response of the recording medium which has not already been compensated for by the assumption in 1.2. The second degradation is the addition to $\mathscr{D}$ of recording noise $n = n(x,y)$.

The only type of degradation treated here is that which can usefully be approximated by convolution of p with a psf $h = h(x,y)$. The ideal degraded image $b = b(x,y)$ is defined by

$$b = p \otimes h. \tag{2.1}$$

In certain situations the degradation is effectively 1-D, which means that the psf assumes the form

$$h(x,y) = h(x)\ \delta(y) \tag{2.2}$$

where $\delta(\cdot)$ denotes the Dirac delta function (a 1-D degradation is always taken to exhibit its variation in the x-direction).

In order to express f in a form convenient for later manipulation, the truncated ideal image, $t = t(x,y)$, is introduced. It is the part of b within Γ:

$$\left.\begin{aligned} t &= b, && \gamma \in \Gamma \\ &= 0, && \gamma \in \Gamma. \end{aligned}\right\} \tag{2.3}$$

If f were identical with b, p could be restored perfectly by inverse filtering (see 1.4.2.1). In chapter 5, it is shown how to estimate b from f, on the assumption that the latter is the best available estimate of t. Consequently, all

parts of f, other than t, are referred to as the contamination c = c(x,y), so that

$$f = t + c \tag{2.4}$$

where

$$c = \mathscr{D}\{p\} - b + n , \qquad \gamma \in \Gamma . \tag{2.5}$$

From now on we take Φ in (1.28) to be the ratio of $|C|$ to $|P|$, where the capital English letters denote the Fourier transform (FT) as defined in (1.8).

The classes $\mathscr{S}$ and $\mathscr{G}$ of degraded images, defined by

$$f \in \mathscr{S} \quad \text{if} \quad t \equiv b, \tag{2.6}$$

$$f \in \mathscr{G} \quad \text{if} \quad t \not\equiv b, \tag{2.7}$$

are now introduced.

The origin of image space coordinates is always placed at some suitable "central point" within Γ. The point spread frame Υ is defined as that part of $\mathscr{J}$ throughout which h spreads the central point of the original image. As is obvious on physical grounds, the size of Υ is always effectively finite for any real-world psf. Note that it is an essential characteristic of SI degradations that each point of p is spread out within a "domain of influence" of $\mathscr{J}$ congruent to Υ but with its centre transformed to the said point. For convenience, Υ is sometimes taken to be somewhat larger than is strictly necessary - in the sense that Υ need only be large enough to span all points of $\mathscr{J}$ at which h is above some threshold. This is done so that Υ can have a convenient shape (e.g. rectangular or circular).

The inverse FT of $\bar{H}$ is called the "inverse psf" and is denoted by $\bar{h}$. The inverse point spread frame $\bar{T}$ is defined in terms of $\bar{h}$, in the same way that T is defined in terms of h.

Two further frames are required for later use. For defining these frames it is convenient to introduce a frame function $\mathscr{L}$, itself defined by

$$\left.\begin{aligned} \mathscr{L}\{\Theta\} &= 1, \quad \gamma \in \Theta \\ &= 0, \quad \gamma \notin \Theta \end{aligned}\right\} \tag{2.8}$$

where Θ is a finite domain of $\mathscr{I}$. The frame Ω is the largest domain of $\mathscr{I}$ that can be occupied by an original image, the domains of influence of all of whose points intersect Γ i.e.

$$\left.\begin{aligned} \mathscr{L}\{\Gamma\} \otimes \mathscr{L}\{T\} &\neq 0, \quad \gamma \in \Omega \\ &= 0, \quad \gamma \notin \Omega. \end{aligned}\right\} \tag{2.9}$$

The frame Λ is the domain of $\mathscr{I}$ occupied by b when p fills Ω i.e.

$$\left.\begin{aligned} \mathscr{L}\{\Omega\} \otimes \mathscr{L}\{T\} &\neq 0, \quad \gamma \in \Lambda \\ &= 0, \quad \gamma \notin \Lambda. \end{aligned}\right\} \tag{2.10}$$

It should be noted that

$$\Gamma \subset \Omega \subset \Lambda. \tag{2.11}$$

It is necessary for later use to define the "extent" of a frame in an arbitrary direction. Let ξ and η denote

Cartesian coordinates inclined at an arbitrary angle to the Cartesian coordinates x and y in $\mathscr{I}$. Now consider two straight lines in $\mathscr{I}$ parallel to the η-axis. The extent of Θ in the ξ-direction is defined as the greatest separation of the lines such that they are tangent to, but do not intersect, Θ.

It is often convenient for frames to be rectangular. The sides of the rectangles can usually be chosen to be parallel to the x and y-axes, in which case a frame Θ is said to be rectangular ($\alpha \times \beta$) if its extents in the x and y-directions are α and β respectively. It is said that Θ is centred-rectangular if its centre coincides with the origin of the x,y-coordinates. Fig. 2.2 shows Ω and Λ for two types of psf, with Γ centred rectangular ($A \times A'$). Fig. 2.2(a) relates to a 1-D degradation, as in (2.2), and Υ is centred-rectangular ($L \times 0$), where L is the extent of h in the x-direction. In Fig. 2.2(b), Υ is centred-rectangular ($L \times L'$).

Unless stated otherwise the psf is assumed to be normalised such that

$$H(0,0) = \iint_{\Upsilon} h(x,y)\,dxdy = 1 \qquad (2.12)$$

for 2-D degradations, and

$$H(0) = \int_{-L/2}^{L/2} h(x)\,dx = 1 \qquad (2.13)$$

for 1-D degradations. The interval $-L/2 \leqslant x \leqslant L/2$ is always used for Υ when the degradation is 1-D (e.g. refer to

Fig. 2.2(a)). It is convenient to define the "resultant psf" $s = s(x,y)$ by

$$s = h \otimes \bar{h}, \tag{2.14}$$

where it is assumed that s and h are similarly normalised. It should be noted that s ideally reduces to the Dirac delta function $\delta(x,y)$ when $\bar{h}$ corresponds to direct inverse filtering as in (1.27).

2.2 CIRCULAR PROPERTY OF CONVOLUTION USING THE FFT

As is mentioned in 1.4.2, the FFT, being a computer algorithm, acts on sampled data. For this reason it calculates a Fourier series rather than a true FT. In other words the sampled data is considered to be taken from a unit cell of a function which is periodically repeated throughout $\mathscr{I}$. For this reason the FFT is said to be "circular" (Silverman 1973). This may be illustrated as follows. Consider the FFT of f in Γ (centred-rectangular $(A \times A')$ as in Fig. 2.2(b)). Let Θ_f be a frame centred-rectangular $(\frac{1}{A} \times \frac{1}{A'})$ in $\mathscr{F}$. The FFT of f in Γ does not give F in Θ_f. Instead it gives the Fourier transform, F_c, of

$$f_c = f(x,y) \otimes \mathrm{comb}(x/A) \otimes \mathrm{comb}(y/A') \tag{2.15}$$

evaluated in Θ_f, where

$$\mathrm{comb}(x/A) = \sum_{n=-\infty}^{\infty} \delta(x-nA). \tag{2.16}$$

Taking the FT of (2.15) gives,

$$F_c(u,v) = F(u,v)\,\mathrm{comb}(Au)\,\mathrm{comb}(A'v), \tag{2.17}$$

which only has value on a rectangular grid of sampling points in Θ_f. Thus, f is considered in (2.15) to be repeated throughout $\mathcal{I}$ within contiguous rectangular domains congruent to Γ, thereby giving rise to f_c.

2.3 PROBLEMS ASSOCIATED WITH CLASS $\mathcal{G}$ IMAGES

Degraded images are arranged into class $\mathcal{A}$ and class $\mathcal{G}$ in 2.1 because in the real-world it is only possible to record things of finite size, i.e. an actual degraded image can only be recorded within a finite Γ. It is usual for a degraded image to occupy a larger region of $\mathcal{I}$ than that occupied by the original, undegraded image. A degraded image which fits completely inside Γ is in class $\mathcal{A}$. A degraded image which is truncated by Γ is in class $\mathcal{G}$.

As stated in 1.4.4.2, standard inverse filtering techniques work well with images of class $\mathcal{A}$. They often give unacceptable results with images of class $\mathcal{G}$, however. The following two reasons for this apply when Γ is rectangular, which is the case of most practical interest because the FFT can be used (with a consequent increase in computational efficiency):

(i) Using the FFT to invert $F\bar{H}$ leads to a $\hat{p}$ existing in Γ; whereas p occupies a domain of $\mathcal{I}$ larger than Γ. In fact p can fill Ω. So, as well as being only an estimate of p, $\hat{p}$ is necessarily incomplete.

(ii) The circular property of the FFT discussed in 2.2, introduces no difficulties when $f \in \mathcal{A}$, because f is everywhere zero (apart from contamination) on the perimeter

of Γ. In this case F is given by

$$F = B + C = PH + C. \tag{2.18}$$

When $f \in \mathcal{G}$, however, the circular property of the FFT can have the effect of making f_c discontinuous on the perimeter of Γ, thereby introducing spurious frequency components into F so that it is no longer given by (2.18). After multiplying by $\bar{H}$ and taking the inverse FT, these spurious components of F introduce artefacts which tend to mask the true detail in p.

Multiplying F by $\bar{H}$ in $\mathcal{F}$ is equivalent to convolving f with $\bar{h}$ in $\mathcal{I}$. Huang et al (1971) argue that if the extent of $\bar{T}$ was considerably less than the extent of Γ, in all directions, then $\hat{p}$ would be likely to represent p faithfully throughout Γ, except close to its perimeter. This would mean that p could be recovered throughout most of Γ by Wiener filtering.

We define (see chapter 1 and 2.1) h, c and b such that they occupy finite domains of $\mathcal{I}$, implying that H, C and B are entire (integral) functions of exponential type, which necessarily possess analytical continuations for all finite, complex values of u and v (Bates 1969). H and Φ may be expressed as functions of polar coordinates

$$H(u,v) = H(\rho;\phi) \quad ; \quad \Phi(u,v) = \Phi(\rho;\phi), \tag{2.19}$$

so that $X = X(\rho;\phi)$ may be defined as

$$X = H H^{*} + \Phi^{2}. \tag{2.20}$$

Since H and Φ are entire functions (of exponential type) of ρ, for any ϕ in the range $0 \leqslant \phi < 2\pi$, it follows that X is such a function. Consequently, X is zero at a denumerable infinity of points in the complex ρ-plane (Bates 1969). On referring to (1.28), it is therefore seen that $\overline{H}$ must possess complex poles, so that the area of $\overline{T}$ is actually infinite, although it is often effectively finite because $\overline{h}$ can be negligible outside a finite domain of $\mathscr{I}$. However, the extent of $\overline{T}$ must be much larger than the extent of T in at least one direction. To illustrate this, the form of $\overline{h}$ for three different constant values of Φ is shown in Fig. 2.3 for a particular 1-D degradation characterised by the h shown in Fig. 2.3(a). As Φ increases, $\overline{h}$ decreases markedly in parts of the range $-\frac{A}{2} < x < \frac{A}{2}$, but it is by no means all concentrated close to $x = 0$. This example confirms other computational experience, which suggests strongly that few psf satisfy the conditions necessary for the aforementioned conjecture by Huang et al (1971) to be useful. It is possible, however, to truncate $\overline{h}$ to a satisfactory length, as is done by Arguello et al (1972). This results in a restoration of inferior quality as is shown in chapter 6.

2.4 SOME PROBLEMS ASSOCIATED WITH OPTICAL IMAGE RESTORATION

Most of the problems associated with optical image restoration using a coherent system have already been discussed in 1.5. The problem of class $\mathscr{G}$ images is mentioned in 1.5 and the discussion in the previous section applies in

this case as well as in the digital case. In addition to the class $\mathcal{G}$ problem, there is another aspect of coherent optical systems which appears to have been ignored in publications on image restoration. The intensity distribution over a cross-section of a light beam emitted by a laser is not uniform, but falls off approximately as e^{-wr^2} where r is the radial distance from the centre of the beam and w is a constant.

Because of the lack of published discussion of this problem, it is treated in some detail here with reference to Fig. 1.2. Let $z = z(x,y)$ be the amplitude distribution of the coherent light incident on the front focal plane (a) of lens (b) in Fig. 1.2. Let a film transparency with ideal amplitude transmittivity b be placed in plane (a). It is assumed in this optical context that b has a constant background level outside Λ (see 2.1). The transmitted amplitude distribution is then given by bz so that the distribution incident on the back focal plane (c) is $B \otimes Z$. A film transparency with amplitude transmittivity $\overline{H}$ is placed in plane (c) so that the transmitted distribution from (c) is

$$(B \otimes Z)\overline{H} = (PH \otimes Z)\overline{H}. \qquad (2.21)$$

The final distribution in the back focal plane (e) of (d) is thus

$$\hat{p} = ((p \otimes h)z) \otimes \overline{h}. \qquad (2.22)$$

If, in appropriate units, $z = 1$ then

$$\hat{p} = p \otimes (h \otimes \overline{h}) = p \otimes s \qquad (2.23)$$

which is the ideal result. However, if z is not constant, the $\hat{p}$ given by (2.22) is a restored image with an SV resultant psf, $s = s(x,y,\alpha,\beta)$, given by

$$s(x,y,\alpha,\beta) = (h(x-\alpha,y-\beta)z(x,y)) \otimes \bar{h}(x,y). \qquad (2.24)$$

Note that $s(x,y,\alpha,\beta)$ is the resultant psf at (α,β).
In practice z is not constant, but can be described by

$$z = e^{-w(x^2+y^2)}. \qquad (2.25)$$

The results of a 1-D computer simulation of the effects on s of w being non-zero are now presented.

The psf and the 1-D resultant psf $s = s(x,\alpha)$ are normalised as in (2.13). Ideally $s(x,\alpha) = \delta(\alpha)$. From (2.24) and (2.25)

$$s(x,0) = (h(x)\ e^{-wx^2}) \otimes \bar{h}(x). \qquad (2.26)$$

In Fig. 2.4(a), $s(x,0)$ is shown for three values of w. $\bar{h}$ was obtained using the direct inverse filter given by (1.27). In Fig. 2.4(b), $s(x,0)$ is shown for a single value of w, but for two different $\bar{h}$ obtained by using two different constant values of Φ in the Wiener filter given by (1.28). As $s(x,0)$ in these examples is symmetrical, it is only plotted for $x \geqslant 0$. The psf is taken to be

$$h(x) = \frac{1}{21\Delta}\ \mathrm{rect}\left(\frac{x}{21\Delta}\right), \qquad (2.26)$$

where Δ is the sampling distance. The particular length of h in (2.26) is chosen so that the zeros of H will lie off sampling points. The product hz is plotted in Fig. 2.5 for

the values of w used in Fig. 2.4.

The results shown in Fig. 2.4 illustrate the dependence of s on w and Φ. For a given w, s deteriorates quickly as Φ decreases. For a low Φ, a given w, and a given form of psf, s deteriorates as the extent of T increases. Also, for given w, h, and Φ, $s(x,\alpha)$ becomes worse as α increases. These trends place an upper limit on the accuracy that can be obtained by performing optical image restoration of a class $\mathcal{S}$ degraded image.

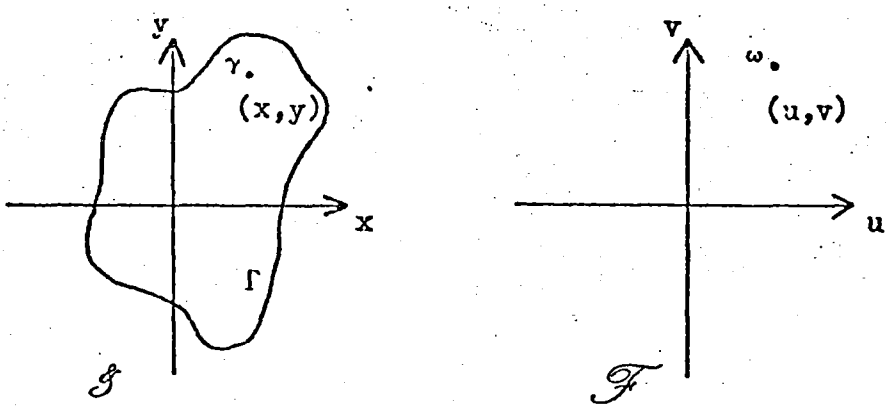


Fig.2.1 Image space $\mathcal{S}$ and Fourier space $\mathcal{F}$.

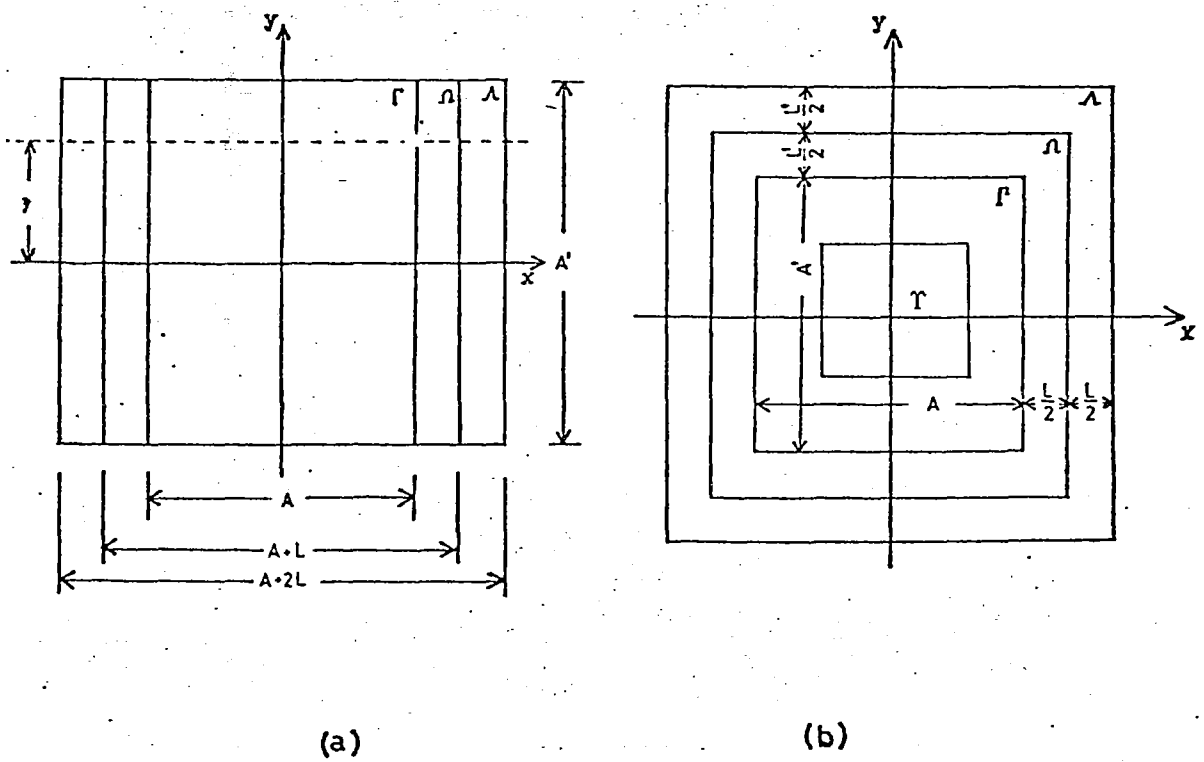


Fig.2.2 Examples of the frames Ω and Λ when Γ is centred-rectangular (A×A').

(a) Υ is centred-rectangular (L×0).

(b) Υ is centred-rectangular (L×L').

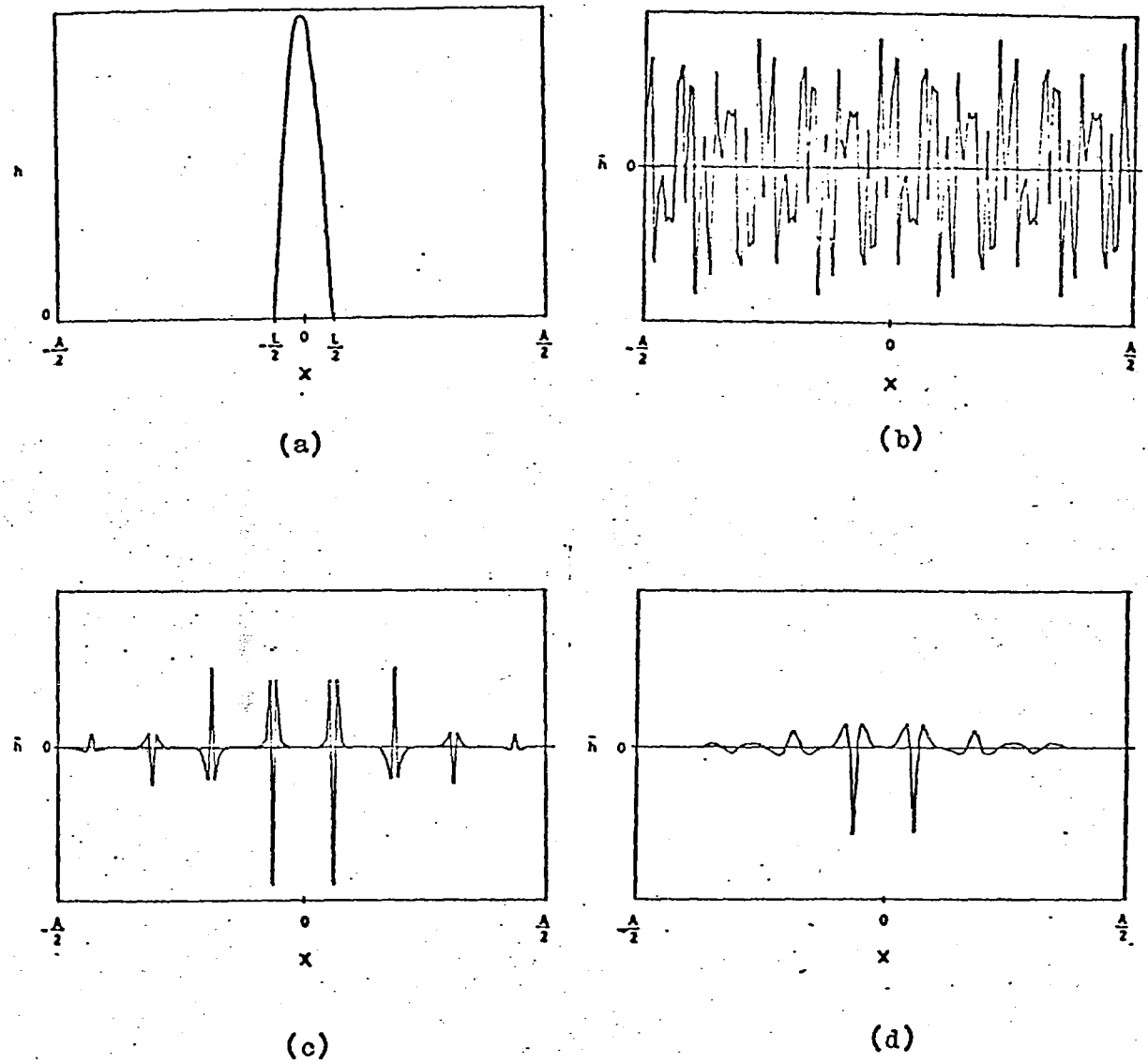


Fig.2.3 Inverse psf for a particular one-dimensional degradation. In each of our one-dimensional examples we take 128 sampling points in the x-direction.

(a) $h(x)$, $-L/2 < x < L/2$.

(b) $\bar{h}(x)$, $W=0$

(c) $\bar{h}(x)$, $W=0.003$

(d) $\bar{h}(x)$, $W=0.03$

$-A/2 < x < A/2$.

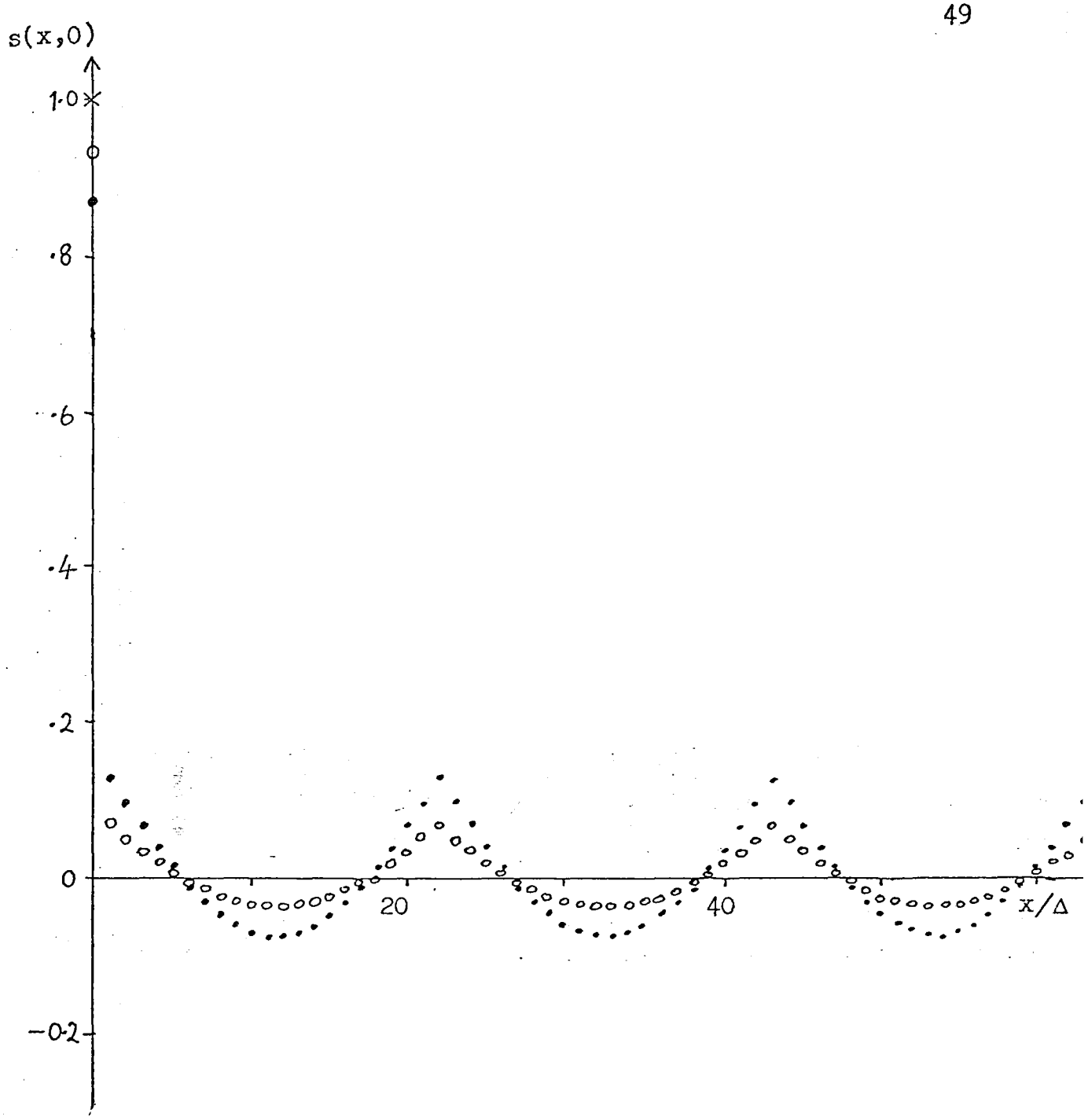


Fig.2.4(a). s(x,0), for $x \geqslant 0$, for three values of w, using the direct inverse filter.

x - - - w = 0,

0 - - - w = 1/512,

• - - - w = 1/256.

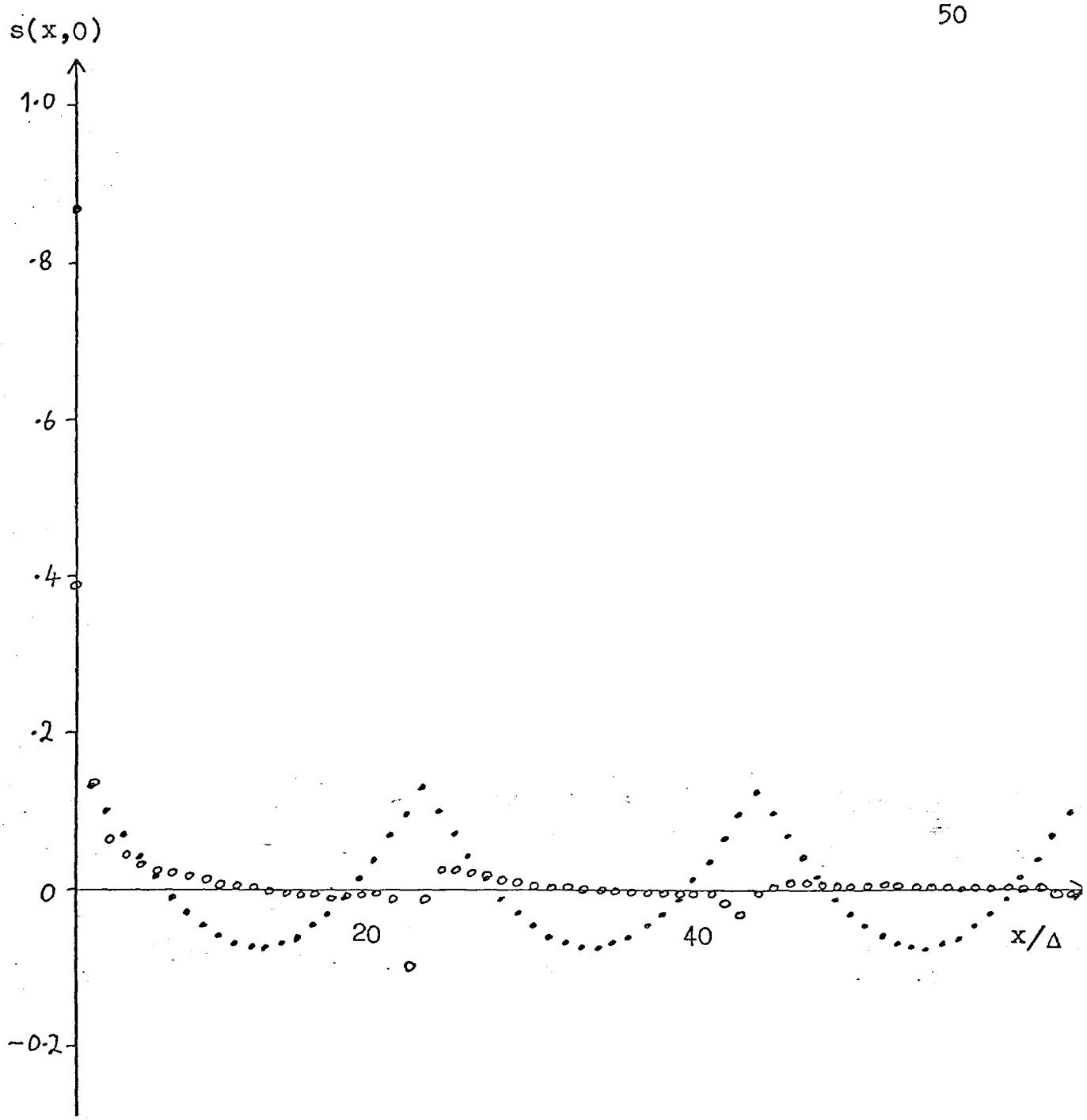


Fig.2.4(b). s(x,0), for $x \geqslant 0$, for $w = 1/256$, but Wiener filtering with

• $\Phi = 0$,

o $\Phi = 0.05$.

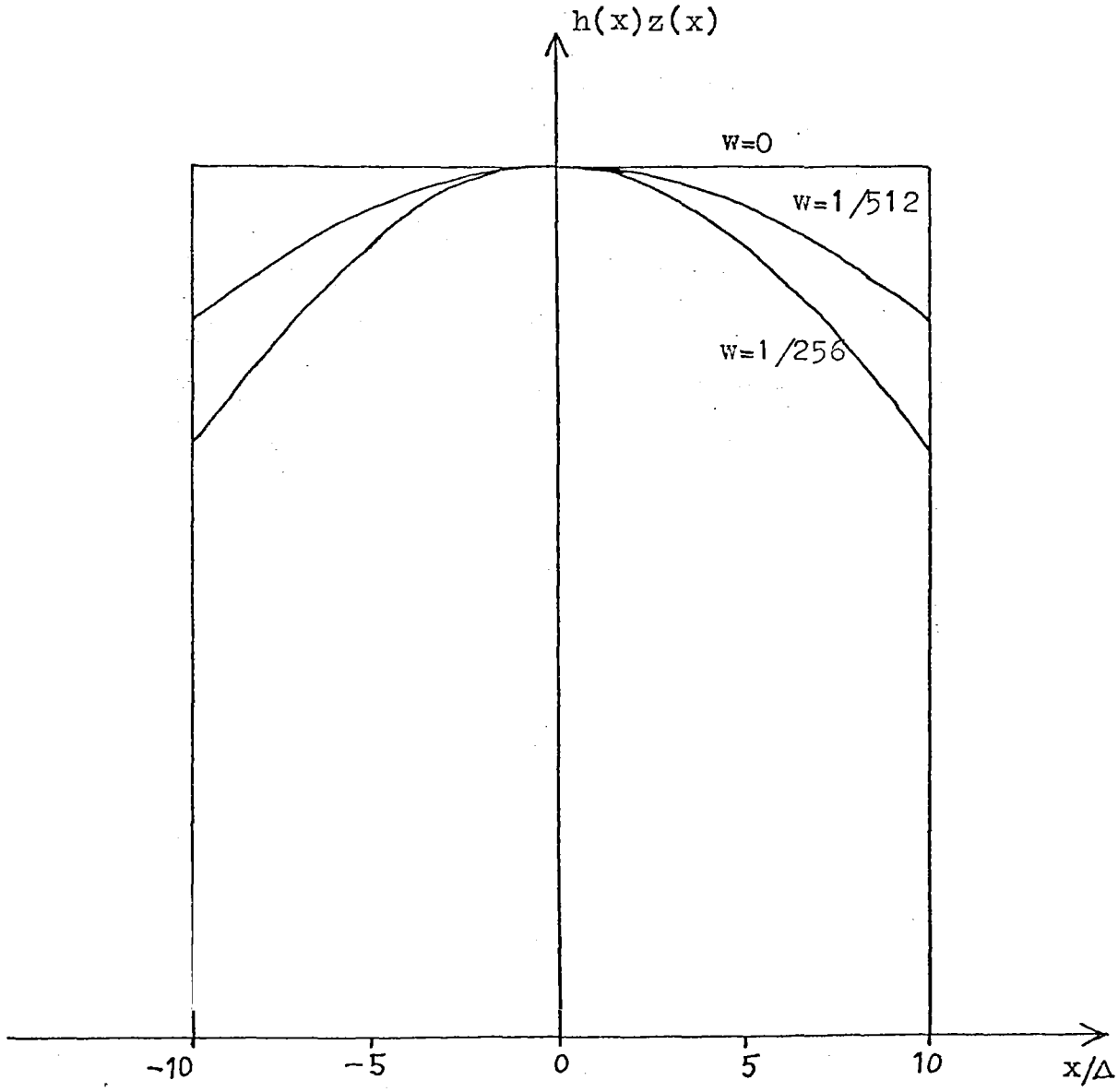


Fig.2.5. The product h(x)z(x) for w = 0,1/512,1/256.

CHAPTER THREE

RESTORATION OF CLASS $\mathscr{A}$ IMAGES

3.1 THE PROBLEM OF THE ZEROS OF H

A detailed discussion of inverse filtering using the FFT is given in 1.4.2.1. In general, inverse filters work well for images of class $\mathscr{A}$. It is only when H is so small that it is below the noise level at sampling points that there is any real difficulty with inverse filtering of class $\mathscr{A}$ images. In this case the direct inverse filter

$$\overline{H} = \frac{1}{H} \tag{3.1}$$

is effectively undefined. The problem can be partly overcome by using the Wiener filter (which is optimum in the MSE sense in the presence of noise), which is defined by

$$\overline{H} = \frac{H^*}{HH^* + \Phi^2} \quad . \tag{3.2}$$

Note that $\overline{H}$ is zero where H is zero because Φ is generally non-zero. The frequency component of the original image at the zero of H is said to be lost (Sondhi 1972) because it is zero in the restored image.

However, if, in the absence of noise, the restored image is obtained using a simple recursive filter (see 1.4.1 and Slepian 1967b) then an ideal restored image is obtained.

In other words there is no information missing at the zeros of H. This implies that Wiener filtering is not a completely satisfactory method of overcoming the problem of the zeros of H. The Wiener filter gives in only the MSE sense the optimum result that can be obtained by inverse filtering. The point here is that inverse filtering alone is not capable of giving the optimum restoration. This suggests that an improved restoration may be possible by modifying the inverse filtering procedure or by using some other operation in addition to inverse filtering.

Bates et al (1974) describe in detail a modified direct inverse filtering procedure which overcomes the problem of the zeros of H for class $\mathcal{A}$ images. This procedure is briefly described in 3.2 and is illustrated with some computer simulations.

In 3.3 it is shown how a recursive operation in addition to inverse filtering can overcome the problem of the zeros of H for a class $\mathcal{A}$ image degraded by uniform linear blur.

3.2 OVERCOMING THE ZEROS OF H BY ANALYTICAL INVERSE FILTERING

In this section and the next the psf is assumed to be due to 1-D motion. For the particular case of class $\mathcal{A}$ images considered in this chapter the full blurred image is contained within the recording frame Γ. Also, a new frame Ψ is defined as the largest domain of $\mathcal{I}$ that can be occupied by an original image, the domains of all of whose points lie

within Γ, i.e.

$$\left.\begin{aligned} \mathscr{L}_{\{\Psi\}} \otimes \mathscr{L}_{\{T\}} &\neq 0, \qquad \gamma \in \Gamma \\ &= 0, \qquad \gamma \notin \Gamma. \end{aligned}\right\} \tag{3.3}$$

The frames Γ and Ψ are illustrated for the 1-D case in Fig. 3.1. The problem is to find $\hat{p}$ in Ψ.

Fourier transforms can be continued analytically into the complex u-plane, and it is straightforward to compute them there (Bates and Napier 1972). Since f exists only within a finite frame Γ, and since the total blurring movement L is finite, it follows that both F(u) and H(u) must be zero at denumerably infinite sets of points in the complex u-plane (Bates and Napier 1972). In this section consideration is given only to the complex zeros of H(u).

Because h(x) is real, H(u) is conjugate symmetric, so that if ζ_m is a complex zero of H(u), where $m \in I_+$, then

$$H(-\zeta_m^*) = H(\zeta_m) = 0, \ \forall m \in I_+. \tag{3.4}$$

Here ζ_m is the mth zero in the right half of the complex u-plane and it is expressable as

$$\zeta_m = \xi_m + i\eta_m, \qquad \xi_m > 0. \tag{3.5}$$

Because it includes some commonly occurring blurring motions, it is convenient to introduce a class Ξ of optical transfer functions defined by

$$\Xi = \{H(u): \xi_m = (m-1+j_m)/\ell, \ \eta_m = \eta; \ m, j_m \in I_+\} \tag{3.6}$$

where ℓ and η are real constants. The essential point is that zeros should occur at points equi-spaced parallel to the real u-axis. However, zeros need not occur at all such points, which is why the term j_m appears in (3.6). It so happens that $H(u) \in \Xi$ when the blurring motion is uniform or sufficiently weakly accelerated that $h(x)$ can be approximated by a truncated exponential; and when the motion is made up of a number of equi-spaced sudden jumps so that the blurred image consists of an overlay of several versions of the original image, with equal shifts between adjacent overlays. For uniform motion blur

$$\left.\begin{aligned} h(x) &= \frac{1}{L}\,\mathrm{rect}(x/L) \\ H(u) &= \mathrm{sinc}\,(Lu) \\ \eta = 0,\ \ell &= L \text{ and } j_m = 1,\ \forall\ m \in I_+. \end{aligned}\right\} \qquad (3.7)$$

When the motion is made up of N equi-spaced sudden jumps then

$$\left.\begin{aligned} h(x) &= \sum_{n=1}^{N} \delta\{x + L/2 - (n-1)L/(N-1)\} \\ H(u) &= \sin\{N\pi Lu/(N-1)\}/\sin\{\pi Lu/(N-1)\} \\ \eta = 0,\ \ell &= L(N-1)/N \text{ and } j_m = 1,\ m < N \end{aligned}\right\} \qquad (3.8)$$

and j_m increases by unity every time m reaches an integer multiple of N (the point being that $H(u) \neq 0$ for $u = kN$, $k \in I_+$).

3.2.1 A Decontamination Procedure

From (2.17) and (3.4) it follows that (neglecting the y-coordinate for a single line of an image)

$$C(\zeta_m) = F(\zeta_m), \quad \forall m \in I_+, \tag{3.9}$$

which suggests that a part of the contamination of the final recorded image may be removed. Let

$$c(x) = c_q(x) + c_r(x) \tag{3.10}$$

where c_r is the removable part of the contamination, and c_q is defined through its FT as

$$C_q(\zeta_m) = 0, \quad \forall m \in I_+. \tag{3.11}$$

The FT of c_r may be expressed in the form

$$C_r(u) = \sum_{n=1}^{\infty} \left[\sigma_n \operatorname{sinc}\{A(u-\zeta_n)\} + \sigma_n^* \operatorname{sinc}\{A(u+\zeta_n^*)\} \right] \tag{3.12}$$

which is quite general because the σ_n are arbitrary. However, it is now postulated that

$$C_r(\zeta_m) = \sigma_m, \quad \forall m \in I_+, \tag{3.13}$$

and a decontaminated blurred image is defined by

$$f_D(x) = f(x) - c_r(x). \tag{3.14}$$

The contamination of the class $\mathcal{A}$ image exists only within Γ. Since the Fourier inverse of (3.12) necessarily exists only within Γ, it follows that the only aspect of the

analysis which is inexact in general is the postulate (3.13). If

$$H(u) \in \Xi \quad \text{and} \quad A/\ell \;=\; J \in I_+ \tag{3.15}$$

then (3.13) is itself exact. Even when (3.15) does not apply, it is nevertheless true that

$$F_D(\zeta_m) \;=\; 0, \qquad \forall\, m \in I_+. \tag{3.16}$$

This decontamination procedure is implemented prior to inverse filtering by first taking the FFT of the samples of f where the sampling distance in $\mathscr{F}$ is $\frac{1}{A}$. The zeros of H(u) are found either analytically or by numerical means (Bates and Napier 1972). The corresponding σ_m are then found from (3.13). This enables C_r to be calculated from (3.12) so that f_D is then obtained from (3.14). If a ζ_m lies on a sampling point, the effect of the decontamination procedure is to force $F(\zeta_m)$ to zero.

In general, for a given A, the larger L is the greater the portion of the contamination which can be removed by the above decontamination procedure. It follows from (3.16) that F_D/H is a well behaved function of u. This allows inverse filtering to be carried out analytically when $H(u) \in \Xi$.

3.2.2 Analytical Inverse Filtering

From now on it is assumed for simplicity that the complex zeros of H(u) are real, i.e. that $\eta = 0$ in (3.6). This assumption is not restrictive because a degraded image in which $\eta = \eta^1 \neq 0$ may be transformed to one in which $\eta = 0$ by premultiplying by $\exp(-2\pi\eta^1 x)$. The restoration is then

carried out with $\eta = 0$, and the final restored image is obtained by multiplying by $\exp(2\pi\eta' x)$. This is discussed by Bates et al (1974) and Kohler and Mandel (1973).

In 1-D the Fourier coefficients F_k obtained from the FFT of f are defined by

$$f(x) = \frac{1}{A} \sum_{k=-K}^{K} F_k \exp(-i2k\pi x/A) \tag{3.17}$$

where $F_{-k}^{*} = F_k$ and $k \in I_+$ such that the F_k are negligible for $|k| > K$.

The Fourier inverse of (3.17) is, by the standard sampling theorem,

$$F(u) = \sum_{k=-K}^{K} F_k \operatorname{sinc}(Au-k). \tag{3.18}$$

From (3.9) it follows that, when $u = \zeta_m$ in (3.18), the left side equals $C(\zeta_m)$. It then follows from (3.12) that, when $u = \zeta_m$ in (3.18), all terms on the right side vanish except one, for which $k = k_m$ say, so that

$$F_{k_m} = \sigma_m. \tag{3.19}$$

Consequently, the decontamination procedure reduces merely to setting to zero all F_{k_m} on the right side of (3.18). We then have that

$$\hat{P}(u) = \frac{F_D(u)}{H(u)} = \sum_{k=-K}^{K(k_m)} \frac{F_k \operatorname{sinc}(Au-k)}{H(u)} \tag{3.20}$$

where the superscript (k_m) on the summation sign indicates that all terms for which $k = k_m$ are missing.

When $H \in \Xi$, straightforward contour integration is

always adequate for providing an analytical formula for $g_k(x)$, the Fourier inverse of $\text{sinc}(Au-k)/H(u)$. It transpires that $g_k(x)$ is of the form

$$g_k(x) = \sum_{m=1}^{M_1} X_m \delta(x-x_m) + \sum_{m=1}^{M_2} y_m' \exp(iw_m' a)\,\text{sqn}(a-y_m), \qquad (3.21)$$

where $\text{sqn}(\cdot)$ denotes the signum function defined by

$$\left.\begin{aligned} \text{sqn}(x) &= -1, \quad x < 0 \\ &= 0, \quad x = 0 \\ &= 1, \quad x > 1, \end{aligned}\right\} \qquad (3.22)$$

and the constants M_1, X_m, x_m, M_2, Y_m, w_m' and y_m depend upon k and the details of the blurring motion. For example, in the case of uniform linear blur, inversion of (3.20) yields

$$\hat{p}(x) = \sum_{k=-K}^{K_{(mJ)}} F_k(-1)^k \sum_{n=-(\frac{J-1}{2})}^{\frac{J-1}{2}} \delta(x-nL)$$

$$+\left(\sum_{n=-(\frac{J-1}{2})}^{\frac{J-1}{2}} \delta(x-nL)\right)\left(\sum_{k=-K}^{K_{(mJ)}} F_k(-1)^k \frac{ki\pi}{JL}\,\text{sqn}(x)\exp\left(\frac{-2\pi ikx}{JL}\right)\right) \qquad (3.23)$$

where the extent of f is chosen so that $A = JL$. For a class $\mathcal{A}$ image this can always be done by increasing A. In the absence of noise, (3.23) is an exact expression for p, which is of extent $(A-L)$. The form of the second part of the right side of (3.23) ensures that the restoration can be performed using the FFT to give samples of $\hat{p}(x)$.

By performing the inverse filtering analytically the Fourier coefficients of the original image at the zeros of H have been recovered. In the presence of noise, some windowing of the higher frequency coefficients is required to reduce the effects of that part of the contamination which was not removed by the decontamination procedure.

The above restoration procedure was first described by Bates et al (1974). A similar approach was later independently adopted by Honda et al (1974).

3.2.3 Results of Computer Simulations

The results of some computer simulations of the restoration procedure described in 3.2.2 are now presented. The original image p is stored digitally in an array of 128 x 128 pixels. A line-printer output of p is shown in Fig.3.2(a). Twenty different density levels are obtained by character overprinting (Gough and Bates 1972).

Fig. 3.2(b) is the result of degrading the image in Fig. 3.2(a) by uniform linear blur for which J = 2 (the scale has been halved in the x-direction in Fig. 3.2(b)). Fig. 3.2(c) is the restoration of Fig. 3.2(b) obtained in the absence of contamination. Fig. 3.3 shows results of applying the technique to overlaid images. A double exposure of the image in Fig. 3.3(a), free of contamination and with J = 4, is shown in Fig. 3.3(b), and its restoration is shown in Fig. 3.3(c). Fig. 3.2(c) and Fig. 3.3(c) are both perfect restorations obtained in the absence of contamination. In both cases the zeros ζ_m of H lie on sampling points. The results show that, in the absence of

contamination, if $H \in \Xi$ the problem of the zeros of H can be overcome by inverse filtering analytically as in 3.2.2. The problem of contamination is left until the next section where it is simpler to deal with.

3.3 SIMPLIFIED RESTORATION OF IMAGES DEGRADED BY UNIFORM LINEAR MOTION

In this section, attention is restricted to the problem of restoring a class $\mathcal{A}$ image which has been degraded by uniform linear motion. It is shown how a recursive operation in addition to inverse filtering can overcome the problem of the zeros of H. The result is similar to (3.23) but is expressed more simply. Also the derivation of the result is less complicated and it includes a simple treatment of contamination.

It should be noted here that one possible way of overcoming the problem of the zeros of H is to sample at such a rate that no zeros lie on sampling points. However, sampling points still lie near zeros. In the presence of contamination, information at these sampling points is only partially recovered by inverse filtering. In the case of the Wiener filter this is because Φ in (3.2) is large near a zero of H. Also, the simple decontamination procedure in 3.2.2 is not taken advantage of if sampling points lie off zeros of H. The techniques in 3.2 and 3.3, however, take advantage of the decontamination procedure and also allow the recovery of the spatial frequencies at the zeros of H.

As in 3.2, A is chosen to be large enough so that

$\frac{A}{L} = J \in I_+$. The restored image $\hat{p}$ should lie inside Ψ and be of extent (A-L). The blurred image lies inside Γ, which is of extent A (see Fig. 3.2). Let Γ/Ψ be the range of x defined by

$$\Gamma/\Psi = \{x : x \in \Gamma, x \notin \Gamma\}. \tag{3.24}$$

Ignoring contamination to start with, the blurred image in Γ is denoted by b. The result of using the FFT to directly inverse filter b in Γ is a restored image, $\hat{p}'$ say, which must also exist within Γ. It is assumed that $\hat{P}'(\zeta_m) = 0$ $(\forall m \in I_+)$ as a result of the direct inverse filtering at the zeros of H (see (3.1)). It follows that

$$\left.\begin{aligned} \hat{P}'(u) &= P(u), \quad u \neq \zeta_m, \quad m \in I_+. \\ &= 0, \quad u = \zeta_m \text{ for some } m \in I_+. \end{aligned}\right\} \tag{3.25}$$

From Table 1.1 it is seen that the zeros of H lie at

$$\zeta_m = \frac{m}{L}, \quad m \in I_+, \tag{3.26}$$

i.e. at every Jth sampling point. The sampling distance in $\mathcal{F}$ is $\frac{1}{A} = \frac{1}{JL}$. (3.25) implies that every Jth frequency component must be removed from p to give $\hat{p}'$. In other words, some error image $\hat{p}_e'$, existing in Γ and of period L, has been subtracted from p to give $\hat{p}'$ in Γ:

$$p - \hat{p}_e' = \hat{p}', \quad \gamma \in \Gamma. \tag{3.27}$$

Since p is zero outside Ψ, only $\hat{p}_e'$ exists in Γ/Ψ which is of total extent L. Hence, the error image of period L may be

determined from $\hat{p}'$ in Γ/Ψ:

$$\hat{p}_e' = -\hat{p}', \qquad \gamma \in \Gamma/\Psi. \tag{3.28}$$

Because $\hat{p}_e'$ is periodic, it is straightforward to find $\hat{p}_e'$ in Γ from (3.28). p is then obtained from

$$p = \hat{p}' - \hat{p}_e', \qquad \gamma \in \Gamma. \tag{3.29}$$

This procedure is illustrated pictorially in Fig. 3.4 for the simple case when p consists of a single delta function.

When the blurred image is contaminated, the image is first restored using the standard Wiener filter. This automatically implements the simplified decontamination procedure in 3.2.2 by setting to zero the contamination at the ζ_m. Next the operation in (3.29) is carried out as before, thus recovering the frequency samples at the ζ_m.

The simplified decontamination procedure removes, on the average, a fraction $\frac{1}{J}$ of the contamination. The remaining contamination is amplified by inverse filtering. Assuming that the resulting noise on $\hat{p}'$ is uncorrelated and of uniform statistics, the noise is then further increased by a factor of $\sqrt{2}$ as a result of the subtraction in (3.29). This is the price paid for regaining the frequency samples of $\hat{P}$ at the zeros of H.

The technique described here can be efficiently implemented with a computer using the FFT. In the absence of contamination it gives identical results to the method at 3.2. However, in the presence of contamination the direct

inverse filtering in (3.20) would need to be replaced by Wiener filtering in order to obtain results comparable to those obtained by the method introduced in this section. In other words (3.20) would become

$$\hat{P}(u) = \frac{F_D(u)H^*(u)}{H(u)H^*(u)+\Phi^2} = \sum_{k=-K}^{K(k_m)} F_k \frac{\text{sinc}(Au-k)H^*(u)}{H(u)H^*(u)+\Phi^2}. \qquad (3.30)$$

The Fourier inverse of (3.30) is difficult to calculate analytically, so that the method of (3.29) gives a more satisfactory treatment of contamination.

3.3.1 Results

Computer lineprinter outputs of restorations obtained using the technique of (3.29) are now presented. Fig. 3.5(b) shows the restoration of the image in Fig. 3.3(a) after it has been degraded in the computer (Fig. 3.5(a)) by a uniform motion blur of known extent. In practice it is necessary to sharpen the results of Wiener filtering at the edges of Ψ before carrying out the subtraction in (3.29). In the examples presented here this was done simply by replacing the two sampled values of $\hat{p}'$ on each side of each edge of Ψ by the values obtained from a linear extrapolation based on the next three sampled values of $\hat{p}'$ further from the appropriate edge of Ψ. Fig. 3.6 shows the restoration of a blurred image to which random additive computer-generated white noise has been added. The standard deviation of the noise is given as a percentage of the maximum intensity of the blurred image. These examples show that the method of (3.29) is a simple efficient way of restoring noisy class $\mathcal{A}$ images degraded by uniform linear blur.

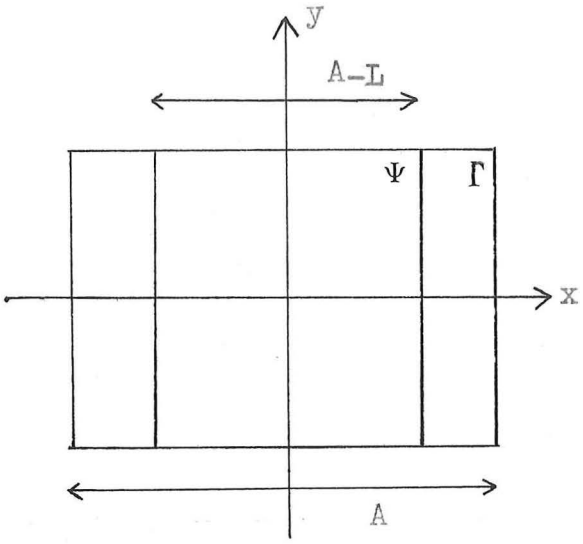
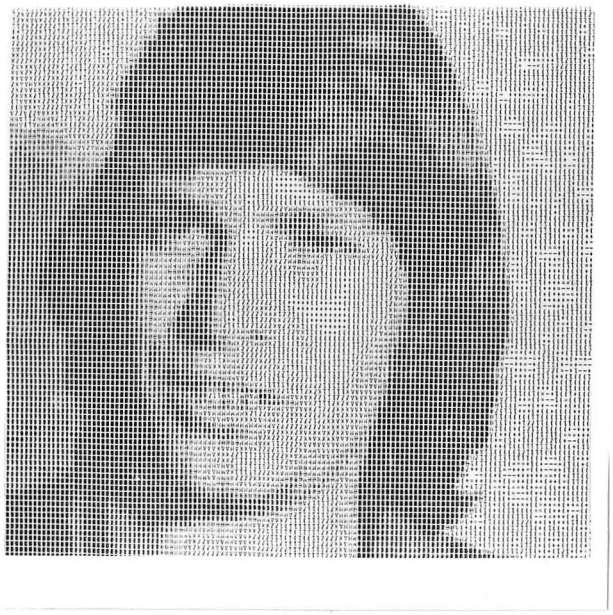
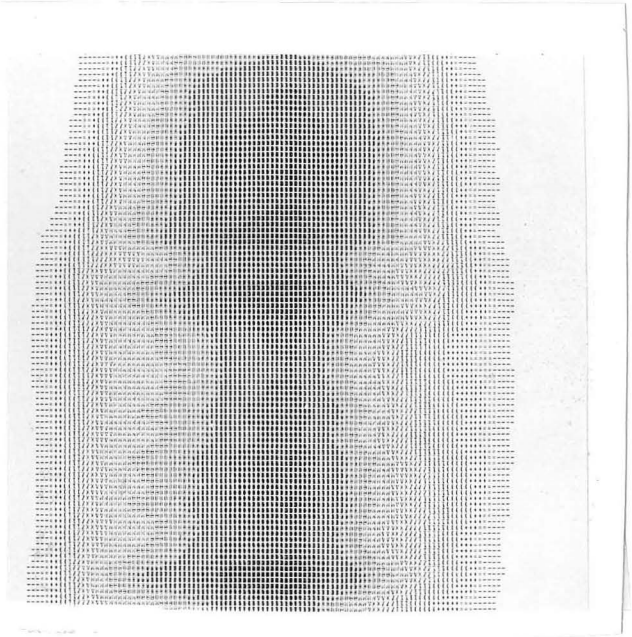


Fig.3.1. The frames Γ, Ψ for a class $\mathcal{S}$ image.



(a)



(b)

(c)

Fig. 3.2. Images in class $\mathcal{S}$ with J = 2 and L = A/2.

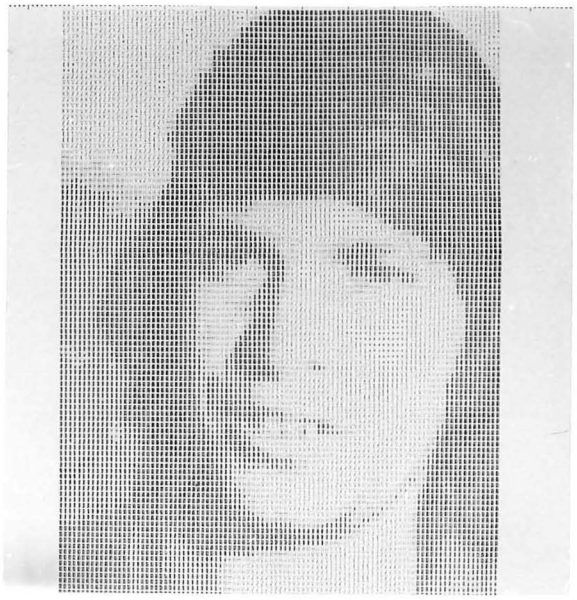
(a) original picture,

(b) recorded picture when blurring motion is uniform and there is no recording noise,

(c) picture restored directly from (b).

(a)

(b)



(c)

Fig. 3.3. Images in class $\mathcal{S}$ with J = 4; the blurring motion consists of sudden jumps with N = 2, $\eta^1 = 0$ and L = A/4.
(a) original image,
(b) recorded picture with no recording noise,
(c) picture restored directly from (b).

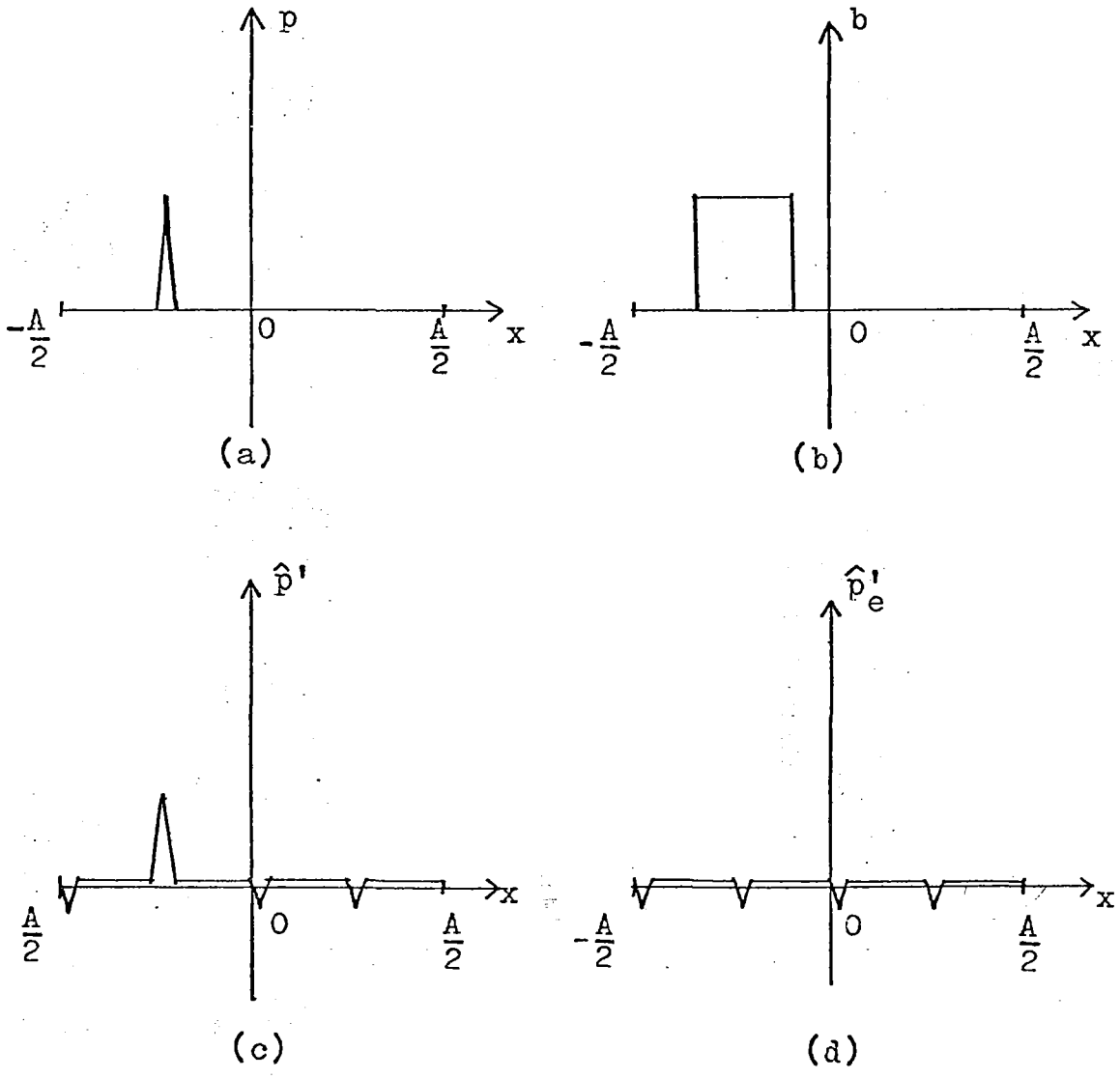
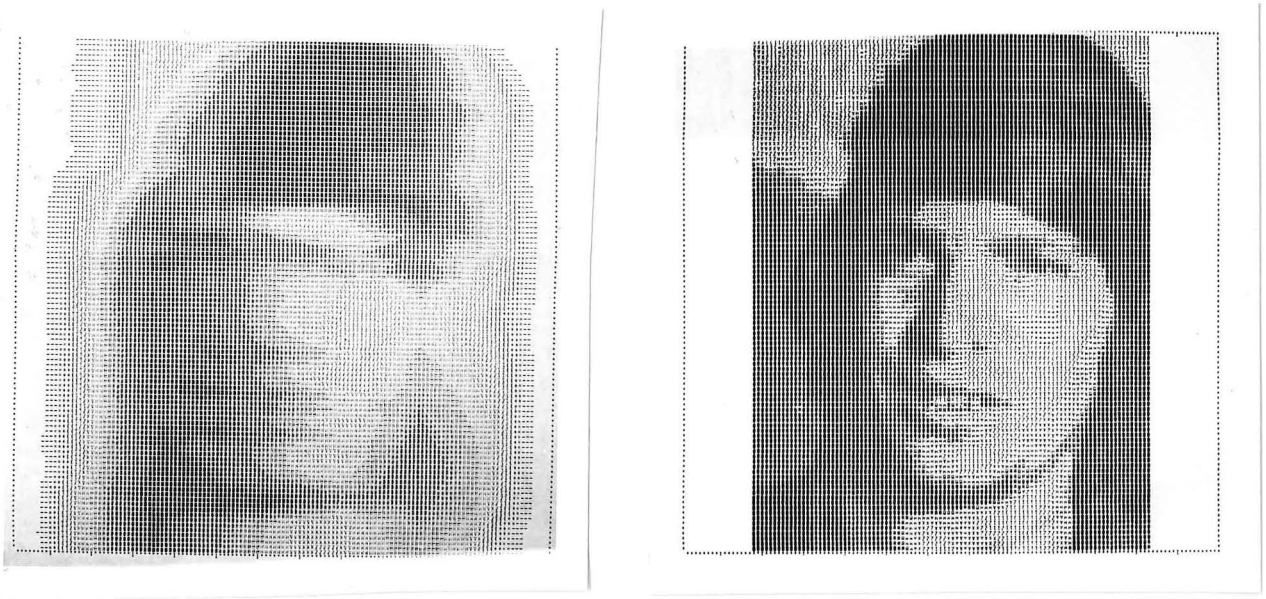


Fig.3.4. Simple illustration of the restoration technique in (3.29). (a) p, (b) b, (c) $\hat{p}'$. (d) $\hat{p}'_e$ which is obtained by periodically repeating $\hat{p}'$ within Γ/Ψ. p, as in (a), is then obtained from (3.29).



(a) (b)

Fig. 3.5. (a) Class $\mathcal{A}$ image resulting from Fig.3.3(a) when the blurring motion is uniform, there is no recording noise, and J = 4; (b) restored image.

Fig. 3.6. restored image corresponding to Fig.3.5(b) in the presence of 3% recording noise.

CHAPTER FOUR

RESTORATION OF RECTANGULAR CLASS $\mathcal{G}$ IMAGES

4.1 INTRODUCTION

The matrix techniques for image restoration discussed in 1.4.4.1 are applicable to class $\mathcal{G}$ images degraded by an SV psf. However, as is illustrated by Mascarenhas and Pratt (1975), such techniques are extremely slow and cumbersome even for images with a small number of pixels (e.g. 64 x 64). Another major disadvantage of the matrix methods of 1.4.4.1, for dealing with class $\mathcal{G}$ images, is that they require the blurred and original images to be sampled at different rates. This is a poor discrete/discrete model of the blurring and leads to instabilities. Practical restoration of the full original image in Ω which contributed to the degraded image in Γ (refer to Fig. 2.2(b)) requires the use of fast algorithms such as the FFT. This, in turn, requires that the degraded image either is SI or is transformable into a form which is SI (Sawchuk 1973). However, as is discussed in 2.3, the FFT can only be used directly to restore an SI degraded image of class $\mathcal{S}$. In this chapter it is shown how to overcome this problem by preprocessing a class $\mathcal{G}$ image into a form which is equivalently of class $\mathcal{S}$. Inverse filtering may then be carried out as efficiently as for class $\mathcal{S}$ images.

4.1.1 Modulation Transfer Function

The concept of the modulation transfer function MTF is widely used in the context of optical imaging. It is defined as the modulus of the optical transfer function H. It is used to describe the attenuation of frequency components of the original image caused by the degradations in the imaging process. This concept of the MTF is valid for degraded images of class $\mathcal{S}$. However, it is not valid in the case of class $\mathcal{G}$ images. This is because a single frequency component in the original image gives rise, not merely to a single frequency component in the degraded image, but to a spread of frequency components caused by the truncation by Γ. The concept of MTF for class $\mathcal{G}$ degraded images does, however, have meaning when the class $\mathcal{G}$ image is preprocessed into an image which is effectively of class $\mathcal{S}$.

4.2 ONE-DIMENSIONAL THEORY

When the degradation is 1-D, (2.2) applies which means that (2.1) can be written as

$$b(x,y) = p(x,y) \otimes h(x) \qquad (4.1)$$

so that it is convenient to process the degraded image along straight lines in $\mathcal{I}$ (on each of which y is a constant). In this section it is shown how to obtain from f, knowing h, an estimate $\hat{b}$ of b, for a particular value η of y, given that Γ is centred-rectangular $(A \times A')$ and Υ is centred-rectangular $(L \times 0)$; where $f \in \mathcal{G}$ and $|\eta| < \frac{A'}{2}$ (see Fig. 2.2(a)).

Two edge extension methods are now introduced. They

are illustrated with an example based on the truncated ideal degraded image shown in Fig. 4.1(a). In this section the main purpose is to explain the details of the methods whereby $\hat{p}$ is recovered from t, and it is convenient to do this without introducing the extra complication of contamination. The example presented in 4.3 demonstrates how the methods behave when recording noise is present.

4.2.1 Full Edge Extension

The curve shown in Fig. 4.1(a) is the intensity of t along the part of the dashed line in Fig. 2.2(a) which lies in Γ. t was obtained by truncating the b shown as the solid curve in Fig. 4.1(b). This b is the convolution of the p shown as the solid curve in Fig. 4.1(c) with the h shown in Fig. 2.3(a).

In practice, recording noise is always present, so it is convenient to think of the t shown in Fig. 4.1(a) is if it were the given data, i.e. f. Since h and therefore L are known, it must be assumed that b has significant value throughout $-\frac{A}{2} - L < x < \frac{A}{2} + L$. So it is desirable to extend t = f along the parts of the dashed line in Fig. 2.2(a) which lie in Λ but outside Γ. It is true that

$$b(-\tfrac{A}{2} - L, \eta) = b(\tfrac{A}{2} + L, \eta) = 0 , \qquad (4.2)$$

$$b(\pm \tfrac{A}{2}, \eta) = f(\pm \tfrac{A}{2}, \eta) \qquad (4.3)$$

and, unless the intensity of the original image $p(x,\eta)$ and/or its derivatives with respect to x are discontinuous within $\frac{A}{2} < |x| < \frac{(A+L)}{2}$ (i.e. in the part of Ω which lies outside Γ)

then $b(x,\eta)$ must be well behaved along the dashed line in Fig. 2.2(a). It is therefore assumed, because it is reasonable to do so and also because there is no simple alternative available, that $b(x,\eta)$ is an analytic function throughout $\frac{A}{2} < x < \frac{A}{2} + L$. So, $f(x,\eta)$ is extrapolated smoothly (i.e. its first few derivatives are matched at $x = \pm \frac{A}{2}$) throughout Λ to obtain the required estimate $\hat{b}(x,\eta)$ of $b(x,\eta)$. The dashed curve in Fig. 4.1(b) was calculated by fitting a polynomial of order 2 to each edge of the t shown in Fig. 4.1(a).

Having obtained $\hat{b}$, an estimate $\hat{p}$ of the original image may now be computed by Wiener filtering. There is no recording noise on the degraded image in Fig. 4.1(a), so that Φ would be arbitrarily small if $\hat{b}$ was an arbitrarily accurate estimate of b. But $\hat{b}$ is not identical with b, so that a finite Φ is required to ensure that $|\hat{B}\overline{H}|$ does not become excessively large for those spatial frequencies for which H is very small (refer to (1.28)). A constant value of 0.03 for Φ proved satisfactory for the examples in this and the following section.

There is some danger in attempting to "optimise" the form of Φ by altering it in some iterative manner until a "best" result is achieved. Consequently, for comparing the results of computer simulations perhaps the only "honest" procedure is to fix on a particular Φ a priori. When the contamination level on f is known then Φ should, of course, be adjusted accordingly. The dashed curve in Fig. 4.1(c) shows the result of using the FFT to compute the inverse FFT

of $\hat{B}\overline{H}$, with $\Phi = 0.03$.

The dashed curve in Fig. 4.1(c) is a poor estimate of the solid curve at the edges. There are two reasons for this. The first is that, because the FFT is used for computational efficiency and economy, the resulting $\hat{p}$ exists throughout Λ, whereas it should only exist throughout Ω. The second reason is that, although $\hat{b}$ is obtained by extrapolating f smoothly, there is no guarantee of the existence of a function (of extent $(A+L)$, such as the p in (4.1), which when convolved with h gives $\hat{b}$.

4.2.2 Overlapped Edge Extension

The estimate of the original image is improved by making use of the circular property of the FFT (already discussed in 2.2). Let the periodic function $e(x,\eta)$ be defined by

$$e(x,\eta) = \sum_{m=-\infty}^{\infty} b(x-m[\underline{A}+\underline{L}], \eta) \qquad (4.4)$$

which is seen to be of period $(A+L)$. Sampling $e(x,\eta)$ within the interval $\frac{-(A+L)}{2} < x < \frac{(A+L)}{2}$ and using the FFT gives its 1-D FT, denoted by $E(u|\eta)$. Because $e(x,\eta)$ is periodic, $E(u|\eta)$ only exists when $u = u_m$, where

$$u_m = \frac{m}{(A+L)}, \; m \in I, \qquad (4.5)$$

where I is the set of integers. Consequently, the restored image, which is obtained by using the FFT to compute the inverse FT of $E(u|\eta)\overline{H}(u)$, is periodic of period $(A+L)$ i.e. $\hat{p}$ now exists in Ω, thereby overcoming the first disadvantage of the full edge extension method of 4.2.1.

Recall that b is of extent (A + 2L), because it exists throughout Λ. However, e is of period (A + L). So the edges of the repeated versions of b overlap in e. Specifically the parts of b along the dashed line in Fig. 2.2(a) (lying in Λ but outside Ω) overlap within the intervals

$$(2m+1)\frac{A}{2} < x < (2m+1)\frac{A}{2} + L, \quad m \in I. \tag{4.6}$$

Consequently, e is called the ideal overlapped degraded image.

The circular property of the FFT ensures the existence of a function, itself periodic of period (A + L), which when convolved with h gives e. This overcomes the second disadvantage of the full edge extension method.

If either $p(-\frac{(A+L)}{2}, \eta)$ or $h(\frac{L}{2})$ is zero and if either $p(\frac{(A+L)}{2}, \eta)$ or $h(-\frac{L}{2})$ is zero then

$$\frac{\partial b(\pm(\frac{A}{2} + L), \eta)}{\partial x} = 0 \tag{4.7}$$

as inspection of (4.1) confirms. Assume that the intensities of all original images fall to zero on the perimeter of Ω, so that (4.7) is always true (note that (4.7) is satisfied however large is $|\frac{\partial p}{\partial \nu}|$ on the perimeter of Ω, where the ν-direction is the normal to the perimeter, provided that p = 0 everywhere on the perimeter). It then follows from inspection of (4.4) that e and $\frac{\partial e}{\partial x}$ are continuous at $x = \pm \frac{A}{2}$. Invoking the assumptions concerning b which are made for the full edge extension method, it is seen that

$$\partial^n e(-\tfrac{(A+L)}{2},\eta)/\partial x^n = \partial^n e(\tfrac{(A+L)}{2},\eta)/\partial x^n, \quad n \in \Gamma_+. \qquad (4.8)$$

An estimate $\hat{e}$ of e is obtained, within the interval $-\frac{(A+L)}{2} < x < \frac{(A+L)}{2}$, by extrapolating f into the intervals $-\frac{(A+L)}{2} < x < -\frac{A}{2}$ and $\frac{A}{2} < x < \frac{(A+L)}{2}$. $\hat{e}$ and its first derivatives are made continuous at $x = \pm \frac{A}{2}$. (4.8) with e replaced by $\hat{e}$ is satisfied for as many values of n as it is felt are warranted by the quality of the data. The solid curve in Fig. 4.2(a) shows e, which is computed from b shown in Fig. 4.1(b). The dashed curve in Fig. 4.2(a) shows $\hat{e}$ obtained (with (4.5) satisfied for n = 1 only) from the t shown in Fig. 4.1(a). The dashed curve in Fig. 4.2(b) is the estimate of the original image obtained by Wiener filtering the $\hat{e}$ shown in Fig. 4.2(a), using $\Phi = 0.03$.

The dashed and solid curves of Fig. 4.2(b) are gratifyingly similar, suggesting that the overlapped edge extension method overcomes the deficiencies of inverse filtering either applied blindly or preceded by the full edge extension method.

4.3 TWO-DIMENSIONAL EDGE EXTENSION THEORY

In this section it is shown how to apply the edge extension method to an image which has suffered a 2-D degradation. In 4-2 it is indicated that the overlapping method is superior to the simple method for 1-D degradations. The same is true in 2-D because the FFT retains its circular property, provided that the function being transformed is defined on a rectangular grid of points. The preprocessing

frame Λ can therefore be dispensed with.

In general, Υ is not rectangular when it covers only that part of $\mathscr{I}$ where h is significant which means that Ω is not, in general, rectangular even if Γ is (refer to (2.9)). It may be possible to choose Γ such that Ω is rectangular, but this would involve a tedious deconvolution process to determine Γ, given Υ. It is more convenient to take Γ to be rectangular and to include in Υ parts of $\mathscr{I}$ where h is negligible, so that Υ itself can be rectangular. It is then straightforward to extrapolate f into Ω.

Γ and Υ are defined to be centred-rectangular $(A \times A')$ and $(L \times L')$ respectively as in Fig. 2.2(b). The ideal overlapped degraded image e(x,y) is defined by

$$e(x,y) = \sum_{m,n=-\infty}^{\infty} b(x-m(A+L), y-n(A'+L')) \qquad (4.9)$$

which is estimated within Ω, by extrapolating f into the part of Ω outside Γ. $\hat{e}$ and its first derivatives (normal to the perimeter) are made continuous on the perimeter of Γ:

$$\left.\begin{aligned} \hat{e}(\pm \tfrac{A}{2}, y) &= f(\pm \tfrac{A}{2}, y) \\ \frac{\partial \hat{e}(\pm A/2, y)}{\partial x} &= \frac{\partial f(\pm A/2, y)}{\partial x} \end{aligned}\right\} \quad |y| < \frac{A'}{2}; \qquad (4.10)$$

$$\left.\begin{aligned} \hat{e}(x, \pm \tfrac{A'}{2}) &= f(x, \pm \tfrac{A'}{2}) \\ \frac{\partial \hat{e}(x, \pm A'/2)}{\partial x} &= \frac{\partial f(x, \pm A'/2)}{\partial x} \end{aligned}\right\} \quad |x| < \frac{A}{2}. \qquad (4.11)$$

It is assumed that b is analytic in the part of Ω outside Γ, so that all derivatives of e exist on the perimeter of Ω. To extrapolate f, as many of these derivates are used as it is felt are warranted by the quality of the data.

Fig. 4.3(a) shows an original image before it suffers a computer-generated blurring which results in the truncated ideal degraded image shown in Fig. 4.3(b). The original image consists of 128 x 128 pixels and is produced by character overprinting on computer line-printer output (Gough and Bates 1972). For this example

$$A' = A, \quad L' = L \text{ and } L = 15\Delta \tag{4.12}$$

where Δ is the sampling distance in both the x and y-directions. The psf is

$$\left.\begin{aligned} h(x,y) &= \frac{4}{\pi L^2}, && (x^2+y^2)^{\frac{1}{2}} \leqslant \frac{L}{2} \\ &= 0, && (x^2+y^2)^{\frac{1}{2}} > \frac{L}{2} \end{aligned}\right\} \tag{4.13}$$

which is a good approximation to the psf of an out-of-focus optical system imaging incoherent radiation (Gennery 1973). It is seen from (4.13) that h is zero in the part of Υ which lies outside the circle of radius $\frac{L}{2}$, centred at the origin of coordinates in $\mathscr{I}$.

To restore the 114 x 114 pixel image shown in Fig. 4.3(b), after recording noise is added to it (so that it becomes f rather than t), f is extrapolated from Γ into Ω in the following way. For a particular y in the range $|y| \leqslant \frac{A}{2}$, f is extrapolated into $-\frac{(A+L)}{2} < x < -\frac{A}{2}$ and

$A/2 < x < (A+L)/2$ using cubic polynomials. Having filled the two domains, $-(A+L)/2 < x < -A/2$ & $|y| < A/2$ and $A/2 < x < (A+L)/2$ & $|y| < A/2$, with the estimate $\hat{e}$, f is then extrapolated into $-(A+L)/2 < y < -A/2$ and $A/2 < y < (A+L)/2$ using cubic polynomials.

Fig. 4.3(c) shows the estimate of the original image which is obtained after the edges of the image shown in Fig. 4.3(b) are extended and then inverse filtered. Fig. 4.3(d) shows the estimate obtained when uniformly distributed recording noise, having an r.m.s. value of 1% of the r.m.s. intensity occurring in b, is added to the image shown in Fig. 4.3(b) before the processing is carried out. The same Φ is used when obtaining both of the restored images because the recording noise associated with the second of these images is less than that theoretically demanded by a value of 0.03 for Φ (a non-zero Φ is always used because it is necessary to compensate for the inaccuracies present in the estimate $\hat{e}$ of e).

In practice it is found that in the presence of appreciable amounts of noise the $\hat{e}$ obtained by extrapolating with cubic polynomials can be poor. This is because noise in f near the perimeter of Γ causes large errors in the estimate of the gradient of f at the perimeter of Γ. These errors decrease the correlation between adjacent lines of $\hat{e}$ in the region of Ω which is outside Γ. The problem may be largely overcome by smoothing f along the two rows or columns closest to each side of Γ before extrapolating f into Ω. This approach yielded the image in Fig. 4.4 which should be compared to Fig. 4.3(d). In the absence of noise,

when f in Fig. 4.3(b) was extrapolated into Ω linearly instead of with cubic polynomials, the result was much worse than that shown in Fig. 4.3(c). This emphasises the value of extrapolating smoothly from Γ into Ω.

4.4 DISCUSSION OF THEORY

Two methods have been presented for extending the edges of degraded images so as to make them more suitable for restoration by inverse filtering. Both methods are simple to implement, since they only involve fitting polynomials of low order to given data. The overlapping edge extension method is the more effective, and it is referred to as the "edge extension method" in the rest of the thesis. The results shown in Fig. 4.3(c) and Fig. 4.4 suggest that the method chosen for extrapolating f from Γ into Ω is relatively uncritical. The method by which the cubic polynomials were fitted for Fig. 4.3(c) is incapable of faithfully reproducing variations of b parallel to the perimeter of Ω, within the part of Ω outside Γ, and yet the restoration is encouraging. What seems to be important about the edge extension method is that it forces $\hat{p}$ to occupy a region of $\mathcal{J}$ of the correct size and it avoids the artefacts which usually appear when images are truncated before being processed.

As the ratio (AA')/(LL') decreases, the relative size of the area, into which f is to be extended, increases. This in turn increases the relative size of the errors due to the edge extension, leading to a poorer restored image.

However, on the other hand, as the ratio $(AA')/LL'$ decreases for a given L and L', the efficiency of the restoration increases. This is because the time taken to perform the FFT of an array of N numbers increases as $N \log_2 N$. Therefore, if a large class $\mathcal{G}$ image is to be restored, a considerable increase in computing speed may be achieved by partitioning the large image into a number of smaller images and then restoring each separately. For example, if a 1024 x 1024 image is restored by partitioning it into 128 x 128 images the processing speed is increased by a factor of 3. The results obtained in this chapter suggest that a value of about 8 for $\frac{A}{L}$ and $\frac{A'}{L'}$ is quite satisfactory.

Another advantage of being able to partition a large class $\mathcal{G}$ degraded image is that a different psf may be used for each smaller image. This means that a large class $\mathcal{G}$ image degraded by an SV psf may be treated as a sum of smaller images each degraded by a different SI psf. This results in a more accurate restoration.

The edge extension method (McDonnell and Bates 1975a) appears to be the first account in the open literature of a general method of restoring the whole of an SI degraded image which is of the class $\mathcal{G}$. The nonrecursive finite filter array methods which are discussed in chapter 6 are only suitable for treating the central parts of such an image. It is appropriate here to mention a method which has recently been independently proposed for dealing with the problem of class $\mathcal{G}$ images. Lewis and Sakrison (1975) and Stockham et al (1975) suggest that before inverse filtering

is carried out, the degraded image f in Γ should be multiplied by a window function which falls off smoothly to zero at the edges of Γ, thereby smoothing the frequency components of F. The effect of this technique is to overcome the effect of the sharp truncation by Γ, and it is thus comparable to the full edge extension method discussed in 4.2.1. The restored image is best in the centre, but the technique is not as good overall as the overlapped edge extension method because it does not constrain the restored image to be of the correct extent. It cannot restore the full class $\mathcal{G}$ image, and the smoothing of F reduces the accuracy of the restoration.

4.5 PRACTICAL EXAMPLES

In this section the use of the edge extension method is illustrated with practical examples. Blurred photographic negatives are scanned with a computer-controlled flying spot scanner (Peters 1973). The samples so obtained are stored in a computer and processed to yield a restored image which is printed either by a standard teletypewriter using character overprinting or by an electrostatic line-printer. In 4.5.1 the steps involved in the implementation of the edge extension method are listed and discussed. A number of practical problems occur in digital image restoration. The success of any restoration, given that the restoration technique is sound, depends largely on how the various practical problems are dealt with. In 4.5.2 the practical problems of psf identification are discussed.

Their theory has already been outlined in 1.3. In 4.5.3 some actual results of using the edge extension method are presented and discussed.

4.5.1 Image Restoration System

The image restoration system developed for this research may be summed up by the following steps:

(1) Develop the blurred negative under controlled conditions so that the gamma of the film (Goodman 1968) can be estimated accurately.

(2) Scan the blurred negative with a computer-controlled flying spot scanner, converting the image to a rectangular array of intensity values (typically 128 x 128 pixels). Also scan a density wedge.

(3) Correct for scanning nonlinearities using the results from the density wedge, and also correct for scanning errors as discussed later in this section.

(4) Adjust the intensity values so that they correspond to a film gamma of unity.

(5) Deduce the psf from the blurred image (described in detail in 4.5.2).

(6) Use cubic polynomials (refer to 4.3) to extrapolate the blurred image (given within Γ) so that it fills Ω. This is the essential step in the edge extension method.

(7) Estimate the noise level in the blurred image and restore it by Wiener filtering, using the fast Fourier transform algorithm (FFT) (Huang et al 1971).

(8) Display the restored image in the most suitable form for the extraction of information by a human being.

Some of the above steps are now discussed in more detail. In my experience an error of 10% in the estimate of the gamma of the film has little effect on the quality of the restoration. The gamma can usually be controlled to this accuracy. In practice, however, a developed blurred negative is often presented for restoration, so that it can be difficult to estimate the gamma accurately. In this case, restoration is attempted for a few estimates of the film gamma, and the gamma finally chosen is the one which yields the result that is judged to be the best. It should be noted here that an error in the estimate of the gamma has the effect of degrading the restored image in the sense that details become less distinct, and less meaningful. It does not have the effect of producing artifacts (i.e. false image elements which might be mistaken for true detail).

The scanning errors referred to in step (3) are due to dust and scratches on the negative, as well as to possible errors introduced by the scanning system itself. Typically, the scanning errors are obvious in the scanned image and may often be largely removed in the following way. For each pixel, the local average (i.e. the average of the intensities of the four closest pixels) is evaluated. If the magnitude of the difference between the local average and the intensity of the pixel being examined is greater than a prespecified amount, the intensity of the pixel is replaced by the local average. Some marks may be too wide to be removed from the scanned image in this way. The required amount of "touching up" has to be assessed by

visual inspection. Image restoration tends to be unstable to all kinds of noise. When large scanning errors are not compensated for, they cause a serious loss of quality in the restored image.

In theory, use of the Wiener filter in step (7) requires knowledge of the ratio of the power spectrum of the original image and the noise on the blurred image. This ratio is a function of spatial frequency and is difficult to estimate accurately. For the examples presented here it has proved satisfactory to assume this ratio to be a constant equal to the reciprocal of the square of the estimated signal to noise ratio.

Step (8) is essentially an image enhancement process which makes the best use of the available image display facilities. The main display device is the standard teletype line printer of the Burroughs B6718 computer in the University of Canterbury's Computer Centre. Twenty different intensity levels are produced by character over-printing. An electrostatic lineprinter sometimes gives superior results, as is discussed in 4.5.3. Particular details of interest in an image may be displayed more clearly by restricting the range of intensity values that are displayed. Further enhancement is possible by techniques such as histogram equalisation (Andrews 1974, Andrews et al 1972)

4.5.2 Psf Identification

The success of image restoration depends strongly upon how accurately the detailed form of the psf can be

identified. Even when the general form of the psf is known (as it is, for example, when the image is assumed to be a photograph taken with the camera out-of-focus, in which case the psf can be approximated closely by a uniform disc (Gennery 1971)) the error in the estimation of the relevant parameter(s) (e.g. the radius of the disc, for out-of-focus blur) needs to be less than 10%, if the restoration is to be visually acceptable. The effects of errors in these estimated parameters are similar to, but significantly more critical than, the effects of an error in the estimated film gamma - refer to step (1) in 4.5.1. Sometimes the detailed form of the psf is given a priori - e.g. it can be deduced from the instrument settings in electron microscopy (Erickson and Klug 1971).

Usually, even the general form of the psf is unknown a priori. Analysis of blurred edges in the available image often permits the general and detailed forms of the psf to be estimated (Huang et al 1971). It is assumed that the edges would have been sharp in the absence of blurring. Suitable edges can only be found by visual inspection. So, the final results depend strongly upon the judgement of the person who is doing the restoration.

In general, it is necessary to recognise a number of edges inclined at varying angles. The intensity of the image is examined along straight lines crossing the edges perpendicularly. The derivative of the intensity along such a line is the line integral (often called "projection") of the psf parallel to the edge. Estimates are improved when

possible by taking several scans across parallel edges. Each projection is thus obtained by averaging a number of scans. A single projection is sufficient to characterise the psf, if it is known (or thought) to be either circularly symmetric (as in out-of-focus blur) or one-dimensional (as in linear-motion blur). Little work has been done on blurring processes which vary arbitrarily in two dimensions, but it is known how to estimate any two-dimensional psf from a finite number of projections (Smith et al 1973).

4.5.3 Examples of Restorations

The restored images presented here are photographs of printouts from either a teletype or an electrostatic line-printer. Except when stated otherwise, the computer output has been displayed by teletype character overprinting.

Each blurred image discussed here is of class $\mathcal{G}$. After estimating the size of T, the size of Γ was chosen such that Ω contained an array of 128 x 128 pixels. This expedited the processing and display of the images.

The first example is an improved version of a previous restoration of a vehicle number plate (McDonnell and Bates 1975b). Fig. 4.5(a) is a properly focussed photograph of a vehicle, showing its rear number plate. Fig. 4.5(b) shows part of the same scene photographed with the camera so heavily out of focus that the number plate is undecipherable (the frame Γ is shown). The part of the degraded image within Γ was digitised into an array of 112 x 112 pixels, which were then corrected, as described in steps (3) and (4) of 4.5.1, to give the sampled blurred image. By analysing

the vertical blurred edges in the sampled blurred image, as described in 4.5.2, it was confirmed that the psf was approximately a uniform disc of diameter 17 ± 2 pixels. In order to carry out successfully the edge extension procedure (step (6) of 4.5.1) from Γ into Ω it is necessary that each side of T consists of an integer number of pixels, and it is convenient if this integer is odd. T was chosen to be a square of side 17 pixels. The blurred image was then extrapolated from Γ into Ω, as required by step (6). Fig. 4.5(c) shows the restoration obtained by Wiener filtering assuming the signal to noise ratio to be 20 (this estimate was obtained by visual inspection of the blurred image).

A few months after producing the restoration shown in Fig. 4.5(c), an electrostatic lineprinter became available. It offered the potentiality for improved displays. So the blurred image shown in Fig. 4.5(b) was re-examined, to see if the estimate of the diameter of the disc representing the psf should be changed. It was found to be closer to 18, rather than 17, pixels. So we then chose T to be a square of side 19 pixels (the psf was set to zero outside the disc of diameter 18 pixels). Fig. 4.5(d) is the restored image, which is a slight improvement on Fig. 4.5(c). Since the intensity levels printed out by the electrostatic line-printer are more evenly-spaced than those printed out by the teletype, it was felt that the former should give a better visual effect, as Fig. 4.5(e) confirms.

The advantages of the electrostatic lineprinter over the teletype are only apparent when it is important to

reproduce varying intensity levels accurately. Fig. 4.6(a) is another out-of-focus photograph of the number plate, but the distortion is appreciably less than in Fig. 4.6(b). The diameter of the psf was estimated to be 15 pixels. The restoration obtained assuming the signal to noise ratio to be 20 is shown in Fig. 4.6(b). The teletype displays the restored numbers quite adequately. It is also worth noting that the reduced distortion present in Fig. 4.6(a) makes it possible to partially restore other detail than the number plate. The outline of the "Land Rover" name-plate is visible in the upper right hand part of Fig. 4.6(b) - compare with Fig. 4.5(a).

The restorations shown in Figs 4.4 and 4.5 are representative of the state-of-the-art for the edge extension technique. Two examples are now presented of situations where the results, while promising and of considerable technical interest, are probably only marginally useful from most practical points of view. However, these examples are important because they indicate clearly the kinds of difficulty which must be overcome if image restoration is to become widely accepted and used.

In Fig. 4.7(a) is shown a single frame of an 8 mm movie film. A face in the background was of interest to the police. Fig. 4.7(b) is an enlarged view of this face, blurred by an unknown psf. The part of Fig. 4.7(b) lying within Γ was scanned with a microdensitometer and then digitised. Using the approach described in 4.5.2, the psf was estimated. It was contained within a rectangular Υ of size 13 x 7 pixels. Fig. 4.7(c) shows the best restoration

obtained, even when use was made of the electrostatic line-printer. The signal to noise ratio was estimated to be 12. The quality of this restoration is poor because the resolution being sought is close to the grain limit of the film. The long, wide mark, visible in Fig. 4(b), sweeping up at 45^{o} to the left from the chin needed to be removed. The loss of detail across the nose and mouth in the restored image (Fig. 4.7(c)) is probably caused by a faulty estimate of the amount of touching up needed before restoration is attempted. This conjecture is supported by our successful restoration of the eye - within the Γ shown in Fig. 4(b) there are no obvious blemishes close to where one would expect the eye to be.

The next example illustrates further difficulties with the removal of extraneous imprints on the image. It also shows what happens when it is attempted to restore detail very close to the edge of the recording frame. A problem which occurs occasionally in forensic science is to recover the original serial numbers stamped on a rifle, after they have been filed off and replaced with a forged set of stamped numbers. One spreads an etching material over the numbers and then photographs the result. The etching is greater the more the underlying crystal structure of the metal has been disturbed by the original stamping. An example of such a photograph is shown in Fig. 4.8(a). This image was digitised and it was attempted to subtract out the letters DU and the number 8 superimposed upon the three right hand blurred numbers in Fig. 4.8(a) - DU8 is part of the forged stamping. The digitised version of the part of

Fig. 4.8(a) lying within Γ is shown in Fig. 4.8(b). The restored image is shown in Fig. 4.8(c), from which it seems that the four original numbers on the right were 3419. This is not clear from Fig. 4.8(a). The left hand number is hardly restored at all, presumably because it is bisected by the edge of Γ in Fig. 4.8(a). The rather fuzzy restorations of the other numbers, which were the best that could be achieved are due to errors introduced by the attempted removal of the superimposed letters DU.

Fig. 4.9(d) shows the restoration of the out-of-focus image in Fig. 4.9(b) of the spoke test chart shown in Fig. 4.9(a). The scanned blurred image is shown in Fig. 4.9(c).

4.5.4 Discussion

The average photograph has more than a million effectively independent pixels stored in it. Because of this it would be very protracted and expensive to attempt *a posteriori* digital processing of large batches of images on a regular basis, unless the pictures were extremely valuable - e.g. the images of Mercury obtained from the Mariner 10 spacecraft (Strom 1974). However, it can often be important to restore part of an image. Since the edge extension technique has been shown to be satisfactory for images of class $\mathcal{G}$, it is possible to concentrate on a small portion of a degraded image and restore the corresponding part of the original scene, provided that important detail is not too close to the edge of the recording frame. The results presented in 4.5 suggest that it is reasonable to

be particularly confident when trying to unmask letters or numbers. In my opinion it is not the restoration method, but the display device which needs to be improved in order to reproduce varying intensity levels more faithfully.

The results suggest that little error seems to be introduced by the basic assumption that the blurring is due to a simple convolution process. This might be because small parts of an image have been dealt with - the smaller the portion of a degraded image that is processed, the more likely is the psf to be effectively invariant throughout that portion. The main difficulty is a purely practical one: more effective methods are required for removing dust marks, scratches and any other blemishes which may mar a photograph which is to be processed.

The CPU time taken for each 2-D image restoration presented in this chapter was close to 80 seconds on the Burroughs B6718 computer at the University of Canterbury.

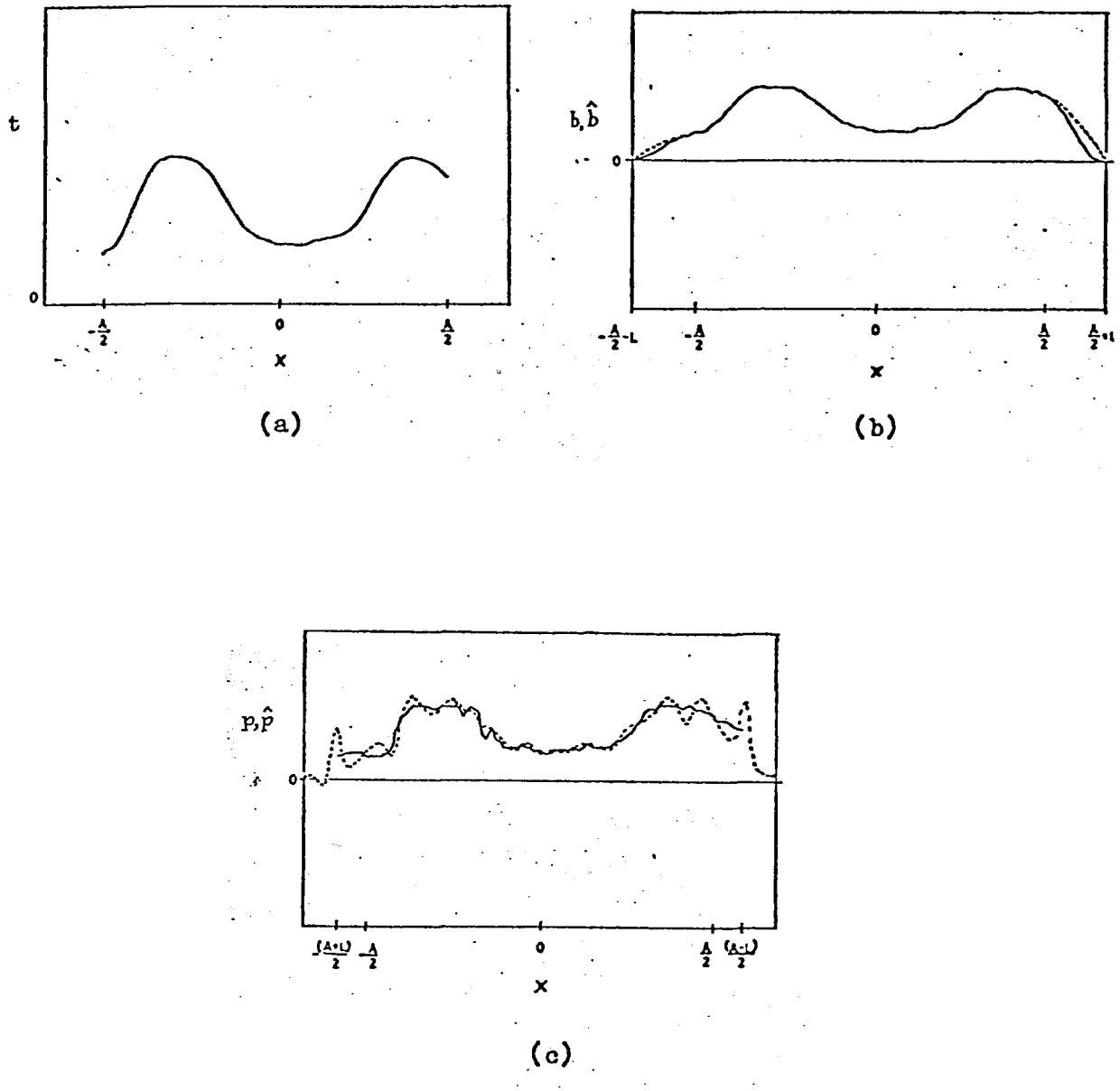


Fig.4.1. An example of full edge extension method.

(a) $t(x,\eta)$, $-A/2 < x < A/2$ i.e. $\gamma \in \Gamma$.

(b) ——— $b(x,\eta)$
- - - - $\hat{b}(x,\eta)$
$-A/2 - L < x < A/2 + L$ i.e. $\gamma \subset \Lambda$.

(c) ——— $p(x,\eta)$
- - - - $\hat{p}(x,\eta)$
$-(A+L)/2 < x < (A+L)/2$ i.e. $\gamma \in \Omega$.

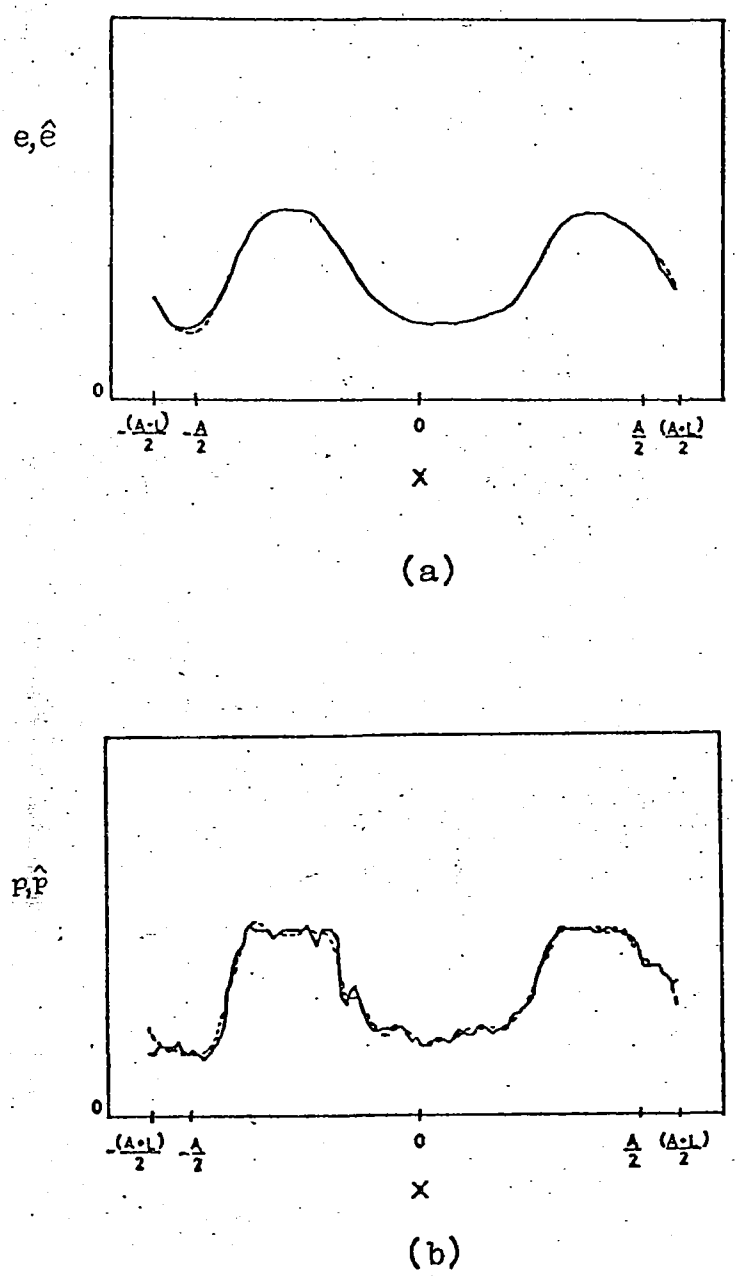


Fig. 4.1. An example of the overlapping edge extension method

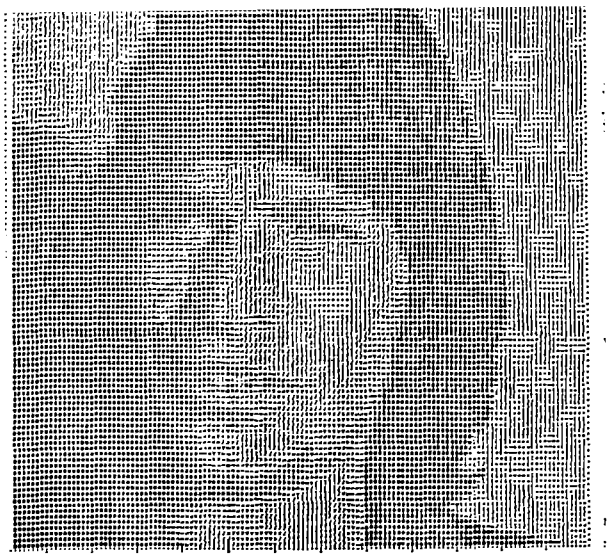
(a) ——— $e(x,\eta)$

- - - - $\hat{e}(x,\eta)$

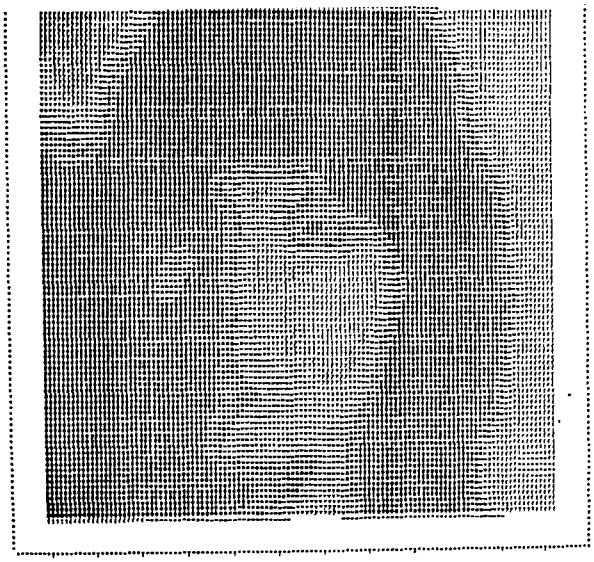
(b) ——— $p(x,\eta)$

. - - - - $\hat{p}(x,\eta)$

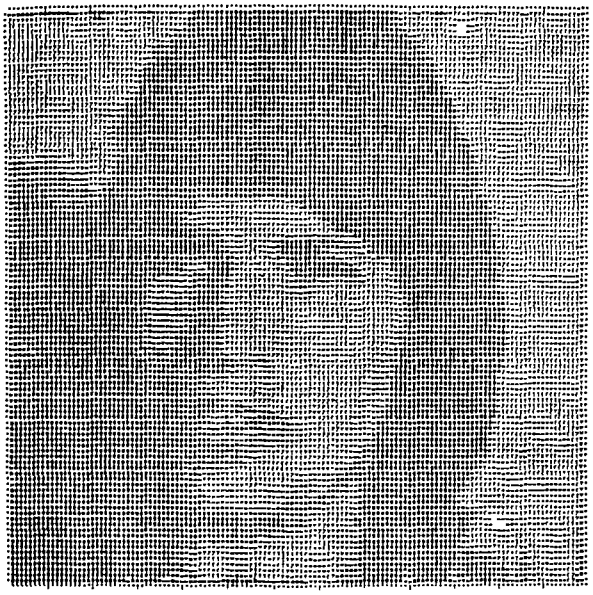
$-(A+L)/2 < x < (A+L)/2$ i.e. $\eta \in \Omega$.



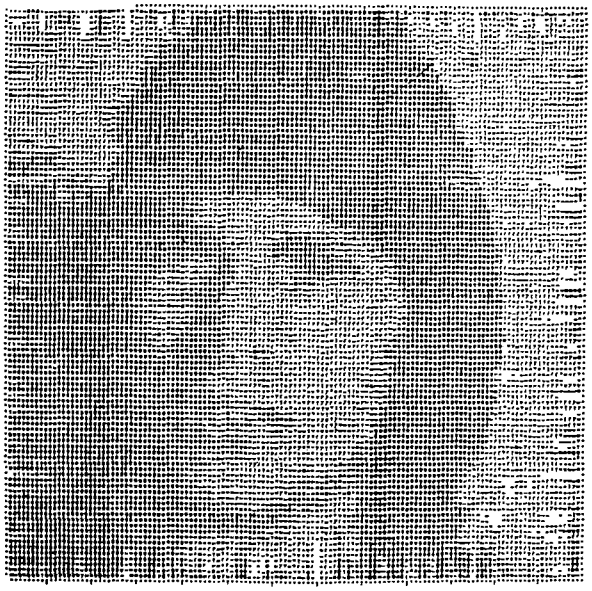
(a)



(b)



(c)



(d)

Fig.4.3. Restoration using edge extension followed by inverse filtering. In Ω there are 128 sampling points in each of the x and y-directions.

(a) $p(x,y)$, $\gamma \in \Omega$.

(b) $t(x,y)$, $\gamma \in \Gamma$.

(c) $\hat{p}(x,y)$, no recording noise on $b(x,y)$

(d) $\hat{p}(x,y)$, 1% recording noise on $b(x,y)$

$\gamma \in \Omega, \Phi = 0.03$.

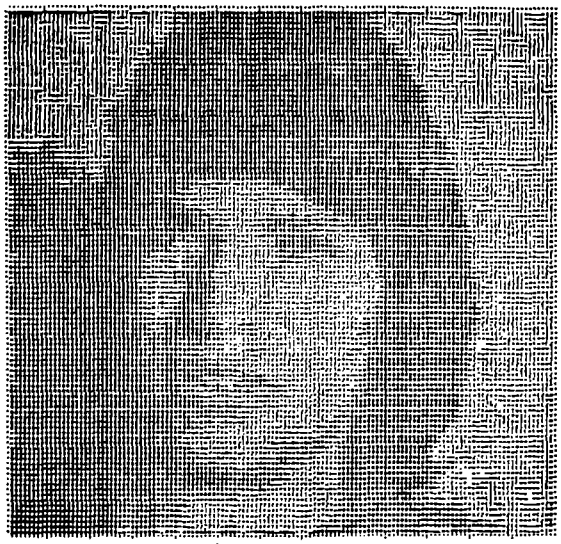
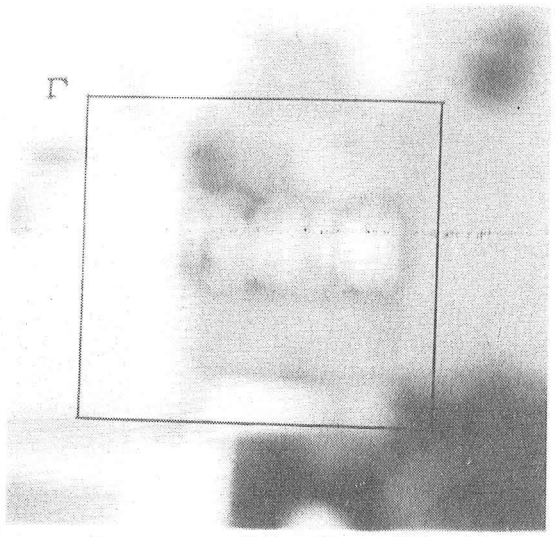


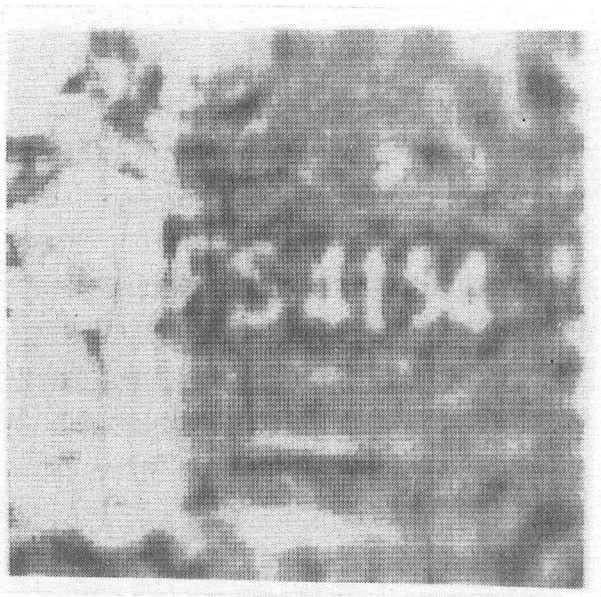
Fig. 4.4. $\hat{p}(x,y)$ as in Fig. 4.3(d) except that $f(x,y)$ has been smoothed around the edges of Γ.

a



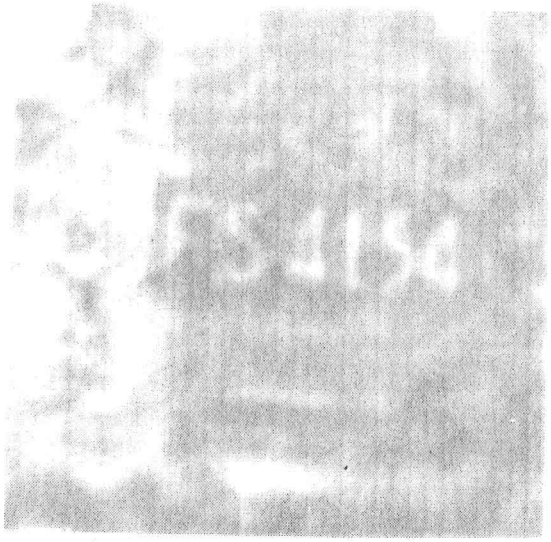
b

C



d

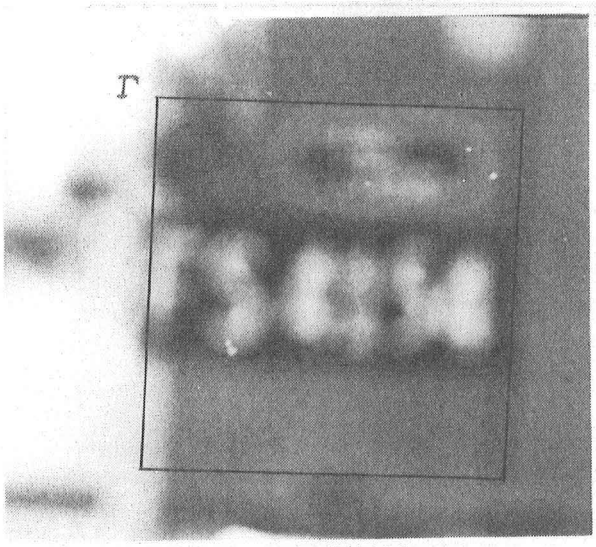
FIG.4.5- Continued on next page.

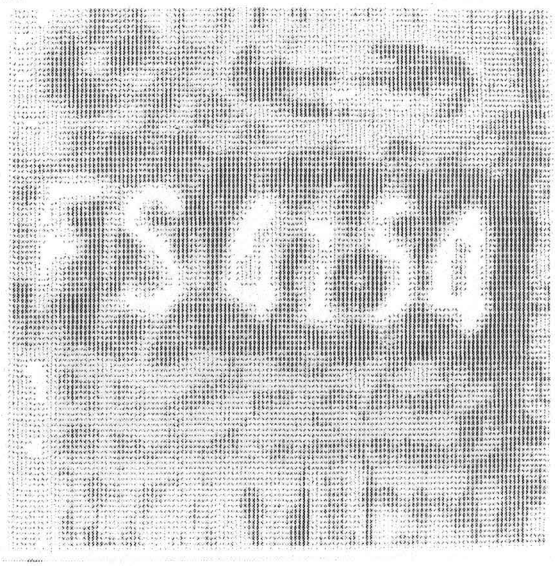


e

FIG. 4.5 - Heavily-blurred photograph of number plate, restored pictures and well-focused photograph

(a) well-focused image (b) out-of-focus image; restored images when r is square of side: (c) 17 pixels, (d) 19 pixels; (e) result of displaying (d) on an electrostatic lineprinter.



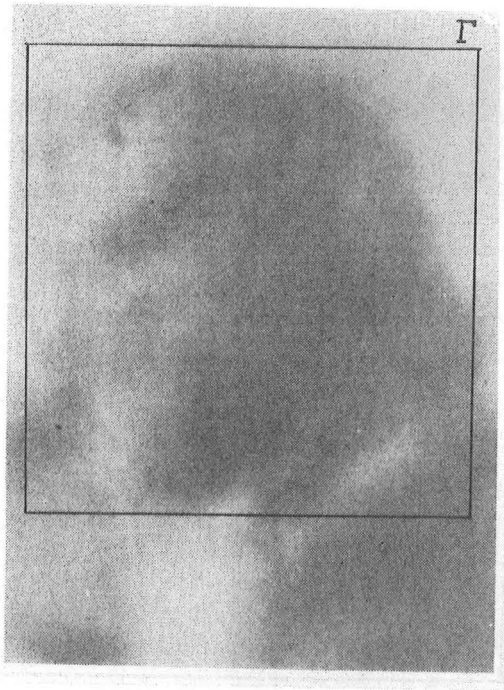


a b

FIG. 4.6 - Moderately-blurred photograph of number plate, and restored picture.

(a) out-of-focus image, (b) restored image.

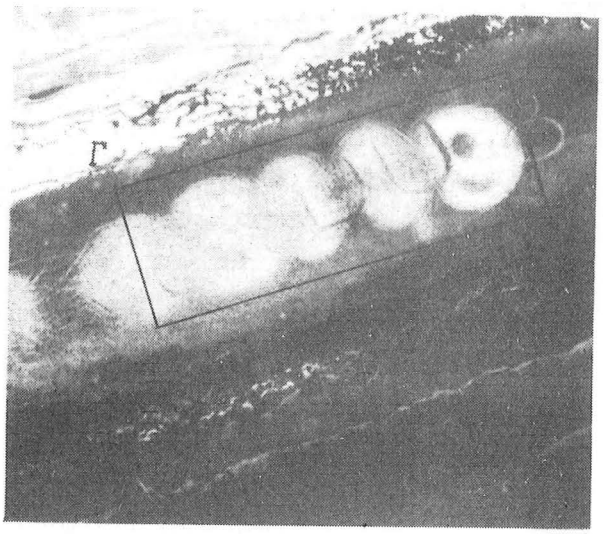
a



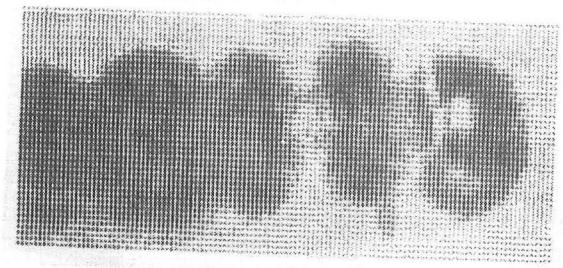
b

c

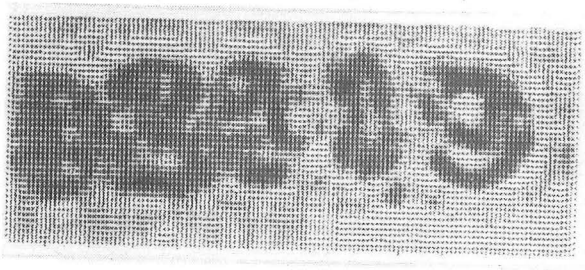
FIG. 4.7 Blurred figure in a street scene and restored picture.
(a) The scene, (b) blurred face (magnified),
(c) restored image (displayed on an electrostatic lineprinter).



a



b



c

FIG.4.8 - Etched serial numbers from a gun, after they have been filed down and over-stamped with forged letters and numbers.
(a) available blurred image, (b) scanned blurred image,
(c) restored image.

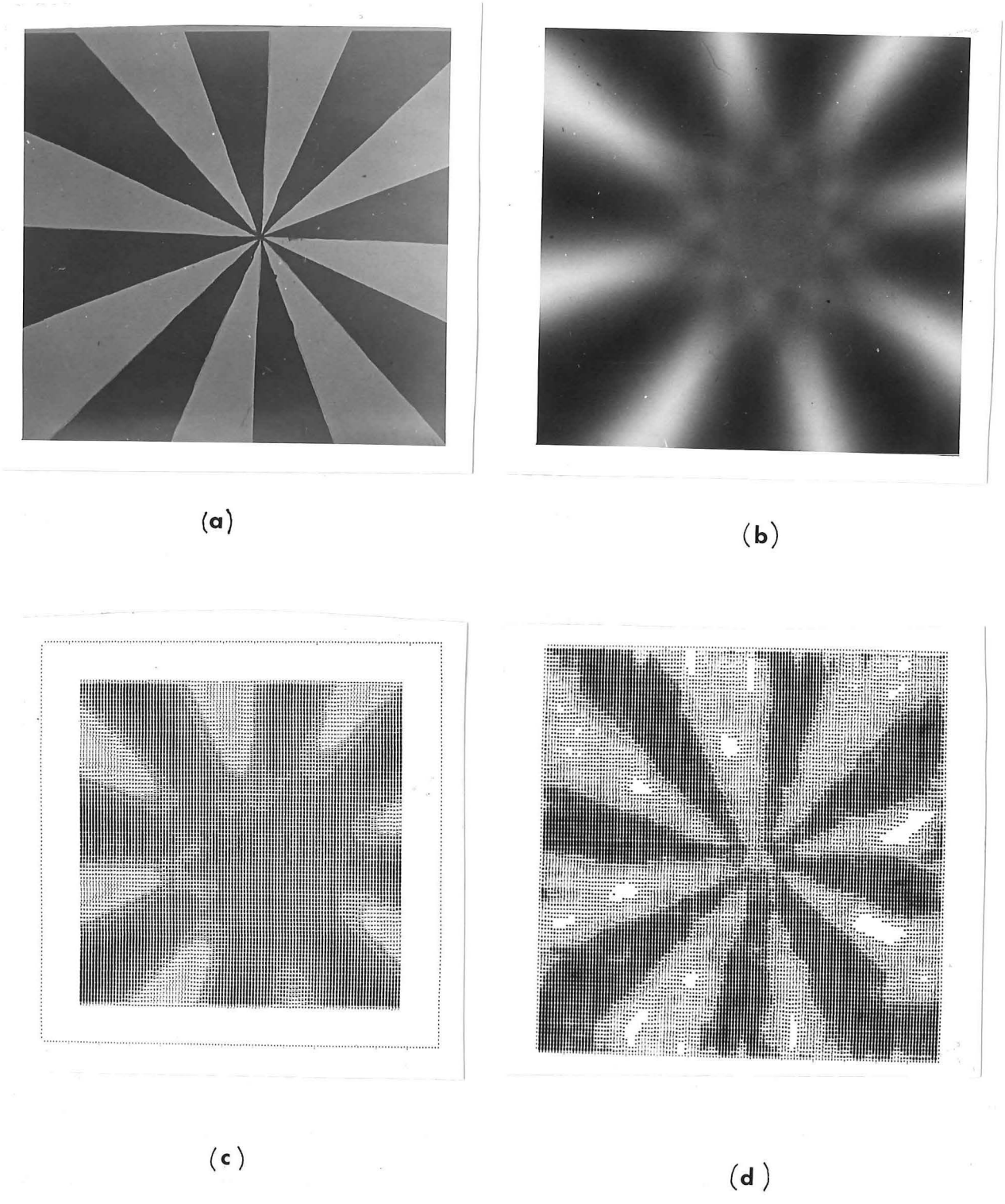


Fig. 4.9. (a) original spoke test chart, (b) out-of-focus blurred image, (c) scanned blurred image, (d) restored image.

CHAPTER FIVE

GENERALISATION OF EDGE EXTENSION METHOD TO INCLUDE ZEROS OF H

5.1 INTRODUCTION

In chapter 4 an edge extension method is introduced by which class $\mathscr{G}$ degraded images may be restored. In that method, f given within Γ (refer to Fig. 2.2) is extended using cubic polynomials to give an approximation $\hat{e}$ to e within Ω. In this chapter it is shown that knowledge of the zeros of H may be used to improve the extrapolation of f from within Γ to give an image which is, or is effectively, of class $\mathscr{S}$. In 5.2 it is shown that for the case of uniform linear blur, and in the absence of noise, e in Ω may be obtained exactly from f in Γ. An improved estimate $\hat{e}$ is also obtained in the presence of noise. In 5.3 the case of circular frames T, Γ, Ω and Λ is considered. It is shown that the zeros of H may be used to obtain, from f in Γ, a good estimate $\hat{b}$ of b in Λ, even in the presence of appreciable amounts of noise. In 5.4 the edge extension method is generalised to show how knowledge of h and f may be used to generate an upper bound $E_{uB}(u,v)$ on $|E(u,v)|$ which may, in turn, be used to obtain an improved estimate $\hat{e}$ of e within Ω.

5.2 RECTANGULAR CLASS $\mathcal{G}$ IMAGES DEGRADED BY UNIFORM LINEAR BLUR

The psf for uniform linear blur, $\frac{1}{L}\,\mathrm{rect}(x/L)$ has the special property that the zeros of H are uniformly spaced (see (3.26)). It is shown in 3.3 that this property may be used to restore perfectly, in the absence of noise, a class $\mathcal{A}$ image degraded by uniform linear blur. The class $\mathcal{G}$ case is considered here, initially in the absence of noise.

Let A be chosen small enough that $(A+L)/L = J \in I_+$. The original image and the overlapped blurred image e (defined by (4.4.)) both lie inside Ω (refer to Fig. 2.2(a)) which is of extent $(A+L)$, whereas f is of extent A. Because FT's are to be carried out using the FFT it is implicitly assumed that both e and f are periodic, with period $(A+L)$. This requires that f be filled out with length L/2 of zeros at each end before the FFT is taken. The zeros of H are given by (3.26):

$$\zeta_m = \frac{m}{L}, \qquad m \in I_+. \tag{5.1}$$

Because e is effectively of class $\mathcal{A}$ (refer to 4.2.2) it follows that

$$E(\zeta_m) = 0, \qquad \forall m \in I_+. \tag{5.2}$$

However, the $F(\zeta_m)$ are, in general, non-zero. Let Ω/Γ be the range of x defined by

$$\Omega/\Gamma = \{x : x \in \Omega,\ x \notin \Gamma\}. \tag{5.3}$$

Define $e' = e'(x)$ by

$$\left.\begin{aligned} e' &\equiv e, && \forall \gamma \in \Omega/\Gamma, \\ &= 0, && \gamma \notin \Omega/\Gamma. \end{aligned}\right\} \tag{5.4}$$

The problem is to find e' from f. The extent of e' is L. From the above definitions it follows that

$$e = e' + f. \tag{5.5}$$

Therefore

$$E = E' + F. \tag{5.6}$$

Hence, from (5.2)

$$E'(\zeta_m) = -F(\zeta_m), \qquad \forall m \in I_+. \tag{5.7}$$

The ζ_m occur at the Nyquist sampling rate (see (5.1)) corresponding to e'. So that, apart from its average value, e' may be found by taking the inverse FFT of the samples of F taken at the zeros of H. For this last FFT, e' is assumed to be of period L. The average value for e' may be found simply by assuming that e in (5.5) is continuous across the edges of Γ. Thus, in the absence of noise, e may be found from f. In the presence of noise a good estimate $\hat{e}$ of e may be found by first windowing the Fourier coefficients of E' given by (5.7) which occur at frequencies at which noise is the dominant component of F, and then proceeding as before.

Fig. 5.1(a) shows a motion blurred image photographed with a standard Pentax 35 mm camera, using ordinary Panatomic-X film. The motion was forced to be linear by

moving the original image on a lathe bed. The blurred image was scanned and then processed as in steps (3) and (4) of 4.5.1, in order to compensate for scanning errors and the gamma of the developed film. This gave the 128 x 112 pixel scanned blurred image shown in Fig. 5.1(b). The extent of the psf was found to be 16 pixels by examining apparent edges in Fig. 5.1(b).

Fig. 5.1(c) shows the restored image obtained by processing Fig. 5.1(b) along 128 lines parallel to the x-direction. Each line was extended from 112 to 128 pixels as described above, windowing the highest frequency components with a cosine bell window. It was then Wiener filtered, letting $\Phi = 0.03$. Fig. 5.1(c) should be compared with the original image shown in Fig. 5.1(d). These results show that the method described in this section is a practical, efficient way of restoring noisy class $\mathcal{G}$ images degraded by uniform linear blur.

5.3 CLASS $\mathcal{G}$ IMAGES IN CIRCULAR FRAMES

In most practical image restoration problems the psf h is two-dimensional. Often h is circularly symmetric (e.g. out-of-focus blur) and generally f is of class $\mathcal{G}$. The image restoration problem is treated here in polar coordinates and a circular Γ of radius R is considered. Let $f = f(r;\theta)$, $h = h(r)$, $b = b(r;\theta)$ and $p = p(r;\theta)$, where the semi-colon is used to indicate the use of polar coordinates. Assume that $h(r) = 0$ for $r > L/2$, and that p exists in a circular frame Ω defined by $r \leq R + L/2$.

It follows that Γ is the disc $r \leq R$, and that b exists in another circular frame Λ defined by $r \leq R+L$ (see Fig. 5.2). Neglect noise initially so that $b = f$ within Γ.

Using the polar coordinates ρ and ϕ as the Fourier transform variables corresponding to r and θ, (1.10) becomes

$$B(\rho;\phi) = H(\rho)\, P(\rho;\phi). \tag{5.8}$$

Approximate $P(\rho;\phi)$ and $B(\rho;\phi)$ by truncated angular Fourier series:

$$P(\rho;\phi) = \sum_{m=-M}^{M} P_m(\rho)\, e^{im\phi} \quad ; \tag{5.9}$$

$$B(\rho;\phi) = \sum_{m=-M}^{M} B_m(\rho)\, e^{im\phi} \quad , \tag{5.10}$$

so that, from (5.8),

$$B_m(\rho) = H(\rho)\, P_m(\rho) \quad , \quad m=-M,\ldots,M. \tag{5.11}$$

The expansion in image space corresponding to (5.10) is

$$b(r;\theta) = \sum_{m=-M}^{M} b_m(r)\, e^{im\theta} \tag{5.12}$$

where b_m and B_m are a Hankel transform pair:

$$B_m(\rho) = 2\pi(-1)^m \int_0^\infty b_m(r)\, J_m(2\pi\rho r) r\, dr \quad ; \tag{5.13}$$

$$b_m(r) = 2\pi(i)^m \int_0^\infty B_m(\rho)\, J_m(2\pi\rho r) \rho\, d\rho \tag{5.14}$$

where $J_m(.)$ denotes the mth order Bessel function of the first kind. Note that

$$b_m(r) = b^*_{-m}(r) \tag{5.15}$$

and

$$B_m(\rho) = B^*_{-m}(\rho). \tag{5.16}$$

Now approximate each $b_m(r)$ as a truncated Fourier-Bessel series

$$b_m(r) = \sum_{n=1}^{N} b_{m,n} J_m\left(\frac{j_{m,n} r}{R+L}\right), \quad m=-M,\ldots,M;\ |r| \leq R+L, \tag{5.17}$$

where $j_{m,n}$ is the nth positive zero of J_m. The form of this series ensures that $b(R+L) = 0$.

Because h is of finite extent, H possesses a denumerably infinite number of zeros $\lambda_k, (k = 1,\ldots,\infty)$. Consider the first N_z of these zeros. It follows from (5.11) that

$$B_m(\lambda_k) = 0, \qquad m=-M,..,M;\ k = 1,\ldots,N_z. \tag{5.18}$$

Using (5.18), (5.13), (5.17) and theorem 8 of Watson (1966, p.134) gives

$$B_m(\lambda_k) = 0 = 2\pi(-1)^m \sum_{n=1}^{N} \left(\frac{b_{m,n}\, j_{m,n}}{\left(\frac{j_{m,n}}{R+L}\right)^2 - (2\pi\lambda_k)^2}\right) J_{m+1}(j_{m,n})\, J_m(2\pi\lambda_k(R+L)), \tag{5.19}$$

$$m=-M,\ldots,M;\ k = 1,\ldots,N_z.$$

Substituting (5.17) into (5.12) gives

$$b(r;\theta) = \sum_{m=-M}^{M} \sum_{n=1}^{N} b_{m,n} J_m\left(\frac{j_{m,n} r}{R+L}\right) e^{im\theta}. \qquad (5.20)$$

If N_f sampled values of f (which equals b in the absence of noise) are given in Γ, then (5.20) gives N_f equations in the $b_{m,n}$. An additional $(2M+1)\,N_z$ equations are provided by (5.19). Taking (5.15) into account, there are $(2M+1)N$ unknowns associated with the $b_{m,n}$. This results in a linear system with $(2M+1)N$ unknowns and $N_f + N_z(2M+1)$ equations. Choosing N such that

$$(2M+1)N < N_f + N_z(2M+1) \qquad (5.21)$$

it is possible to solve for the $b_{m,n}$ using the method of least squares (which is described in 1.3.4). Next, b is found in Λ using (5.20) and the solution for p in Ω is then obtained using a standard deconvolution technique, such as Wiener filtering.

This technique for recovering b throughout Λ is now illustrated with an example in which p is circularly symmetric: i.e. $p = p(r)$ and $b = b(r)$. In Fig. 5.3 are shown h, p and the parts of b which lie inside and outside Γ. Assume that samples of f are given at non-negative integer values of r up to $r = R$. Let $R = 40$, $L = 20$, $M = 0$ and $N_f = 41$. The first four zeros of H are at $\rho = 0.061$, 0.112, 0.162 and 0.212 respectively. Taking $N = 10$ corresponds roughly to a Fourier cutoff for F at $\rho = 0.081$. It follows that for $N = 10$ a choice of N_z greater than 1 is consistent with (5.21).

In Fig. 5.4 is shown the b obtained outside Γ for

$N = 10$ for several values of N_z. Note that the result for $N_z = 4$ is very satisfactory as only ten terms were used in (5.20). In Fig. 5.5 is shown the effect of varying N with $N_z = 4$. For $N > 10$ the results become unstable because of computational errors. In Fig. 5.6 are shown the results obtained for $N = 10$ and $N_z = 4$ when varying levels of uniformly distributed random noise are added to the given samples of f in Γ. The maximum noise level is expressed as a percentage of the maximum value of f. These results show that a satisfactory estimate of b in Λ may be obtained from samples of f in Γ, even in the presence of appreciable noise levels.

5.4 CONSTRAINED EDGE EXTENSION METHOD FOR 1-D CLASS $\mathscr{G}$ IMAGES

The edge extension method introduced in chapter 4 enables f, when it is given within Γ (refer to Fig. 2.2(a)), to be extended using cubic polynomials to give an approximation $\hat{e}$ to e within Ω. But, even in the absence of noise $\hat{e}$ does not equal e. It is simply a good approximation to it. In this section it is shown that for a 1-D class $\mathscr{G}$ degraded image, knowledge of the zeros of H may be used to improve the extrapolation procedure described in 4.2.2; f is extended in such a way that it fills Ω and is equivalently of class $\mathscr{A}$. For simplicity the noise free case is considered initially.

The estimate $\hat{b}$ of b within Λ given by the full edge extension method in 4.2.1 is, in general, inconsistent with

b in the following sense. The FT of (4.1) is

$$B(u) = P(u)\, H(u), \tag{5.22}$$

so that, at a zero ζ_n of H,

$$B(\zeta_n) = 0, \qquad n \in I_+. \tag{5.23}$$

However, in general,

$$\hat{B}(\zeta_n) \neq 0, \qquad n \in I_+, \tag{5.24}$$

so that $\hat{b}$ can be said to be inconsistent with b. This concept of consistency has been introduced and dealt with in detail by Bates et al (1975) and McKinnon et al (1975).

It seems reasonable to suppose that the constraint (5.23) could be used to obtain an improved estimate $\hat{b}$ of b. However, the interest here is in sampled rather than continuous systems. The same idea of consistency is not directly applicable to the case when the FFT, rather than the continuous FT, is used to compute B(u). This is because the zeros ζ_n of H do not, in general, lie on sampling points in $\mathcal{F}$. However, it is possible to formulate a modified consistency condition which is applicable to frequency samples obtained using the FFT.

The ideal overlapped blurred image is defined in (4.4) as

$$e(x) = \sum_{m=-\infty}^{\infty} b(x-m(A+L)) \tag{5.25}$$

$$= b(x) \otimes \sum_{m=-\infty}^{\infty} \delta(x-m(A+L)). \tag{5.26}$$

Taking the FT and using (5.22) gives

$$E(u) = P(u)\, H(u) \sum_{m=-\infty}^{\infty} \delta(u-m/(A+L)) \tag{5.27}$$

which becomes

$$E\left(\frac{m}{A+L}\right) = P\left(\frac{m}{A+L}\right) H\left(\frac{m}{A+L}\right), \quad m \in I, \tag{5.28}$$

at sampling points. This can be rewritten as

$$E_m = P_m H_m, \qquad m \in I. \tag{5.29}$$

Let P_{uBm}, with $m \in I$, be an upper bound on $|P_m|$. From (5.29) it follows that

$$|E_m| \leqslant P_{uBm}|H_m| = E_{uBm}, \quad m \in I. \tag{5.30}$$

This gives a new consistency condition on the samples $\hat{E}_m$ of $\hat{E}$:

$$|\hat{E}_m| \leqslant P_{uBm}|H_m| = E_{uBm}, \quad m \in I. \tag{5.31}$$

In general, the estimate $\hat{e}$ of e in Ω given by the edge extension method does not satisfy (5.31). The problem here is to obtain an estimate $\hat{e}$ of e which does satisfy (5.31). First one obtains an initial estimate $\hat{e}^i$ of e (which will not satisfy (5.31)). Then obtain from $\hat{e}^i$ a satisfactory estimate of the P_{uBm} $(m \in I)$. The E_{uBm} are then found from the P_{uBm} using (5.31). Finally the modulus of each $\hat{E}_m^i$, for which $|E_m^i| > E_{uBm}$, is reduced in turn to E_{uBm} to give an $\hat{E}_m$ satisfying (5.31). For each m, this reduction is achieved by subtracting a complex correction term $C_{n,m}$ from each sample

$\hat{E}^i_n$ of $\hat{E}^i$. The $C_{n,m}$ are chosen in such a way that $\hat{e}^i$ is changed only in Ω/Γ (see 5.3). This is because it is known that e = b within Γ.

The simplest way of obtaining $\hat{e}^i$ is by the edge extension method. It has been found from experience that the following procedure results in a reasonable set of P_{uBm}, given $\hat{e}^i$ and h. The technique relies on the result that |E(u)| generally decreases with u. It is illustrated schematically in Fig. 5.7 using continuous functions although the actual implementation is necessarily discrete. $\hat{e}^i$ approximates the circular convolution of p and h, so that $|\hat{E}^i|$ approximates |E|, the product of |P| and |H|. |P|, |H|, |E| and $|\hat{E}^i|$ are plotted in Fig. 5.7(a)-(c). They are only shown for positive u because they are symmetric about the origin. From $|\hat{E}^i|$ is constructed the minimum monotonically decreasing function $Q_1 = Q_1(u)$ which is everywhere greater than or equal to $|\hat{E}^i|$. This function is illustrated in Fig. 5.7(d). Next Q_1 is Wiener filtered to give $Q_2 = Q_2(u)$ (see Fig. 5.7(e)), assuming in (1.28) a constant value for Φ, which is a rough estimate of the noise level on the given degraded image. From Q_2 is constructed the maximum monotonically decreasing function $Q_3 = Q_3(u)$ which is everywhere less than or equal to Q_2 (see Fig. 5.7(f)). The final estimate $Q_4 = Q_4(u)$ for the upper bound to |E| is then obtained from

$$Q_4 = 1.05\,|H|\,Q_3 + Q_c, \tag{5.32}$$

where the factor 1.05 is included for safety, and Q_c is a small positive constant equal, say, to the maximum value of

$|H|Q_3$ over the highest 10% of the frequency range being considered. The last term Q_c is included in (5.32) so as to reduce the sensitivity of this procedure to high frequency noise. The most important components of $|\hat{E}^i|$ to be corrected are those near the first zero of H. In Fig. 5.7(g), Q_4 and $|\hat{E}|$ are plotted for comparison.

Once the E_{uBm} corresponding to Q_4 in Fig. 5.7(g) have been found, the next step is to set equal to E_{uBm} the modulus of each $\hat{E}^i_m$ for which $|\hat{E}^i_m| > E_{uBm}$. Consider the correction of $|\hat{E}^i_m|$ to E_{uBm}, where $m \in I_+$. The most obvious way of doing this is to subtract a correction term from $\hat{E}^i_m$ and leave the other $\hat{E}^i_n$ $(n \neq m)$ unchanged. Let $C'_{m,m}$ be the apparent complex correction required to be subtracted from $\hat{E}^i_m$ where

$$|C'_{m,m}| = |\hat{E}^i_m| - E_{uBm}, \tag{5.33}$$

and where the phases of $C'_{m,m}$ and $\hat{E}^i_m$ are the same. Because $|E_m| = |E_{-m}|$, the same correction must be applied to both $\hat{E}^i_m$ and $\hat{E}^i_{-m}$, so it is best to consider these two terms together. Another point to note is that the total complex correction subtracted from $\hat{E}^i$ to ensure that $\hat{E}^i_m$ and $\hat{E}^i_{-m}$ satisfy (5.31), must be such that the resulting $\hat{e}$ only differs from $\hat{e}^i$ in Ω/Γ. This cannot be accomplished simply by subtracting $C'_{m,m}$ from $\hat{E}^i_m$ and $C'_{m,-m}$ from $\hat{E}^i_{-m}$, or equivalently by subtracting $c'_m = c'_m(x)$ from $\hat{e}^i$. However, it can be achieved by multiplying c'_m by a window function $w' = w'(x)$ defined by

$$\left.\begin{aligned} w'(x) &= \mu, \quad x \in \Omega/\Gamma; \\ &= 0, \quad x \notin \Omega/\Gamma, \end{aligned}\right\} \tag{5.34}$$

where μ is a real constant. This is equivalent to convolving the frequency correction samples $C'_{m,m}$ and $C'_{m,-m}$ with $W'(u)$ to obtain the $C_{n,m}$ and then correcting $\hat{E}^i$ at each frequency sample. The phases of the $C_{n,m}$ are controlled so that the phases of $\hat{E}^i_m$ and $\hat{E}_m$ are the same. This is achieved simply by multiplying $C'_{m,m}$ and $C'_{m,-m}$ by an appropriate phase factor $e^{i\phi'}$ and $e^{-i\phi'}$ respectively, before convolving with $W(u)$. Each $C_{n,m}$ is given by

$$C_{n,m} = C'_{m,m}\, e^{i\phi'} W'_{n-m} + C'_{m,-m}\, e^{-i\phi'} W'_{n+m} \tag{5.35}$$

where the $W'_j (j \in I)$ are samples of $W'(u)$.

The procedure described in the previous paragraph is carried out for every $m \in I_+$ such that $|\hat{E}^i_m|$ is greater than E_{uBm}. In general, each correction to the mth frequency component disturbs all other frequency components. However, the disturbance is not large. If necessary, the total correction procedure may be iteratively repeated until a satisfactory convergence is obtained. Usually most of the improvement is obtained in the first two or three iterations. From now on, the method of extending edges developed in this section is referred to as the constrained edge extension method.

5.4.1 Results and Discussion

The constrained edge extension method is now

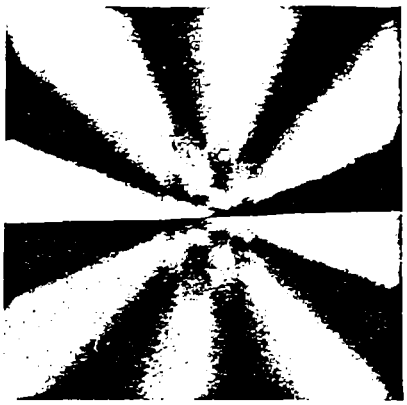
illustrated by the sample in Fig. 5.8. Shown in Fig. 5.8(a) is b given within Γ. This results from the convolution of the p shown in Fig. 5.8(b) with the h in Fig. 5.8(c). In Fig. 5.8(d) are shown $\hat{e}^i$ obtained from the edge extension method and also e for comparison. In Fig. 5.8(e) are shown $|E|$, $|\hat{E}^i|$ and $|\hat{E}|$. Fig. 5.8(f) shows $\hat{e}$ together with e for comparison, after 10 iterations of the procedure to improve the edge extension. The sum of the moduli of the differences between $\hat{e}^i$ and e in Ω/Γ in Fig.5.8(d) is 93. The corresponding error in Fig. 5.8(f) between $\hat{e}$ and e is 41. In this particular example, only slow improvement was obtained by increasing the number of iterations, as shown in Fig. 5.8(g). For comparison, the corresponding reduction in the error between $\hat{e}$ and e is plotted in Fig. 5.8(g) as a function of the number of iterations for the case when $E_{uB} = |E|$. Convergence in this case is much faster.

The above results indicate that in the absence of noise the constrained edge extension method is capable of significantly improving the accuracy of the extension of b in Γ to give $\hat{e}$ in Ω. The method is simple and fairly fast computationally so that it should be practical for use in 1-D signal processing applications.

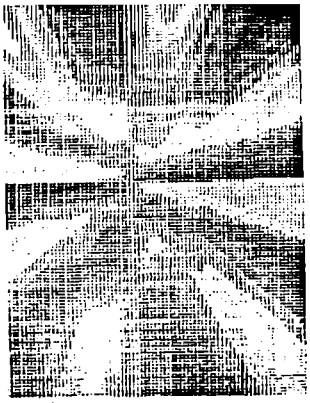
In the presence of noise the method may be used as before except that the $\hat{E}^i_m$ should only be corrected for $m \leq M$, where M is a cut-off value which depends on the noise level. Some noise reduction may be achieved in this case by setting $|\hat{E}^i_m|$ to E_{uBm} when $m > M$ and $|\hat{E}^i_m| > E_{uBm}$. A computer simulation was carried out in which uniformly distributed

random noise of maximum value 3% of the maximum value of b, was added to the b shown in Fig. 5.8(a) to give f. Using the edge extension method the error between $\hat{e}^i$ and $\hat{e}$ in Ω/Γ is 168. The corresponding error after 2 iterations of the constrained edge extension method is 70. The reduction in the error is shown in Fig. 5.8(g) as a function of the number of iterations.

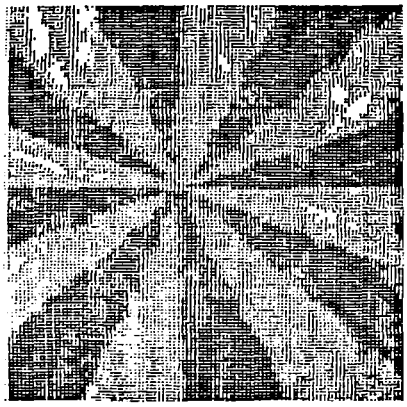
The constrained edge extension method has been shown here to give an estimate $\hat{e}$ of e in Ω which is more accurate than that given by the edge extension method. The greater the number of sampling points over which the edge is to be extended, the greater the advantage of using the constrained edge extension method. In 2-D the method appears to be too slow to be of practical use. However, the technique of determining an upper bound E_{uB} on $|E|$ is still practical and can be used to obtain some noise reduction in the $\hat{e}$ given by the edge extension method.



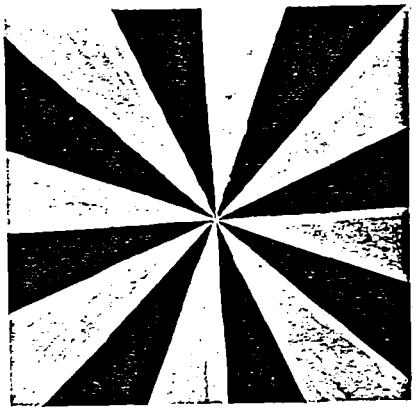
(a)



(b)



(c)



(d)

Fig. 5.1. Images of test chart. (a) Actual blurred photograph of chart. (b) Blurred photograph after scanning. (c) Restored image. (d) Original image.

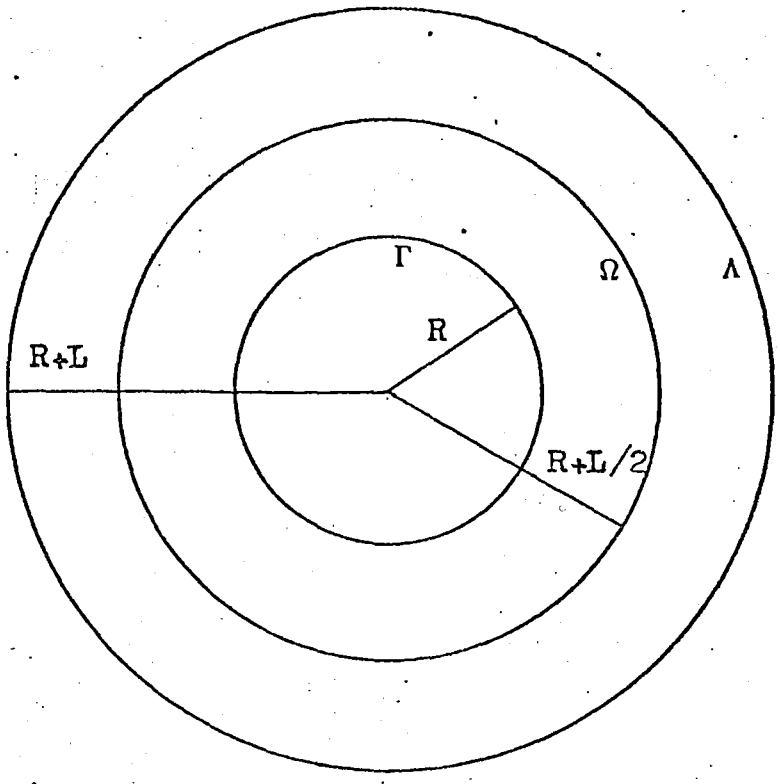


Fig. 5.2 - the circular frames Γ, Ω and Λ.

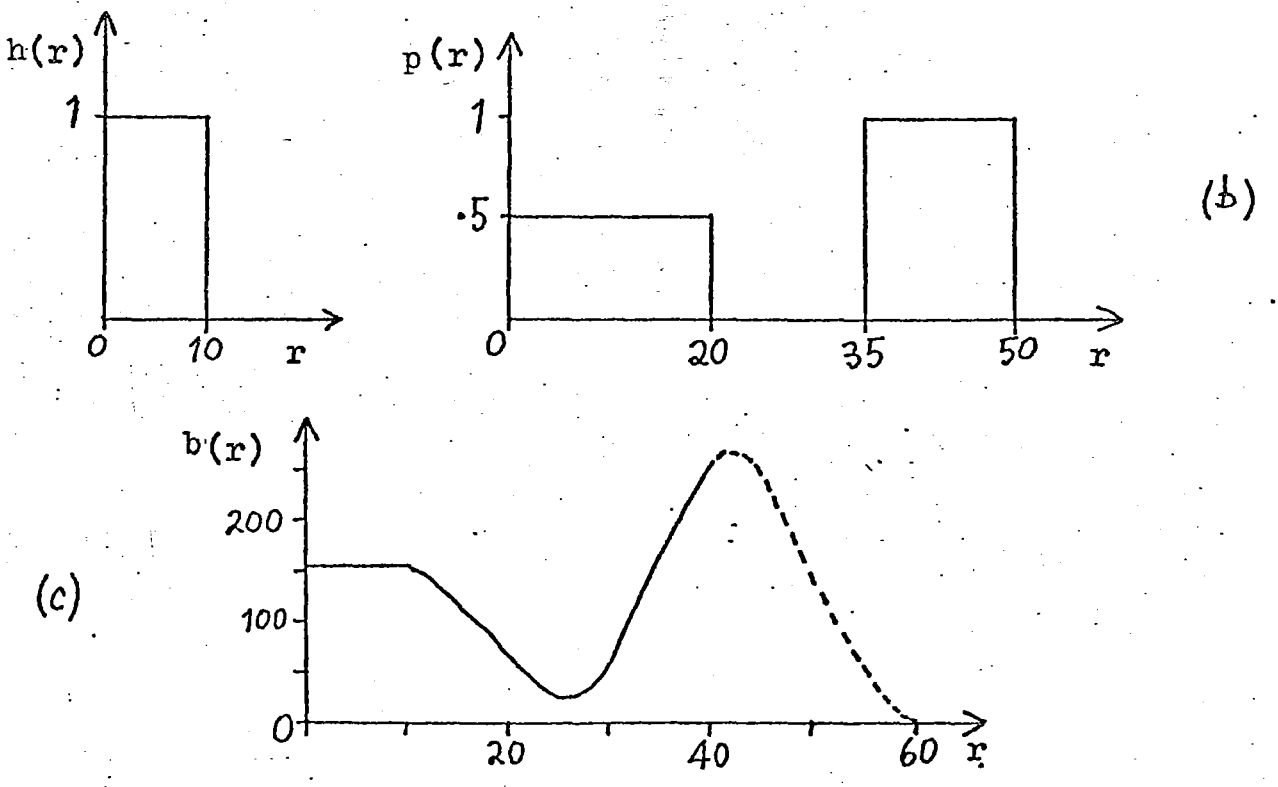


Fig. 5.3. (a) h, (b) p, (c) _____ b in Γ, _ _ _ _ b outside Γ.

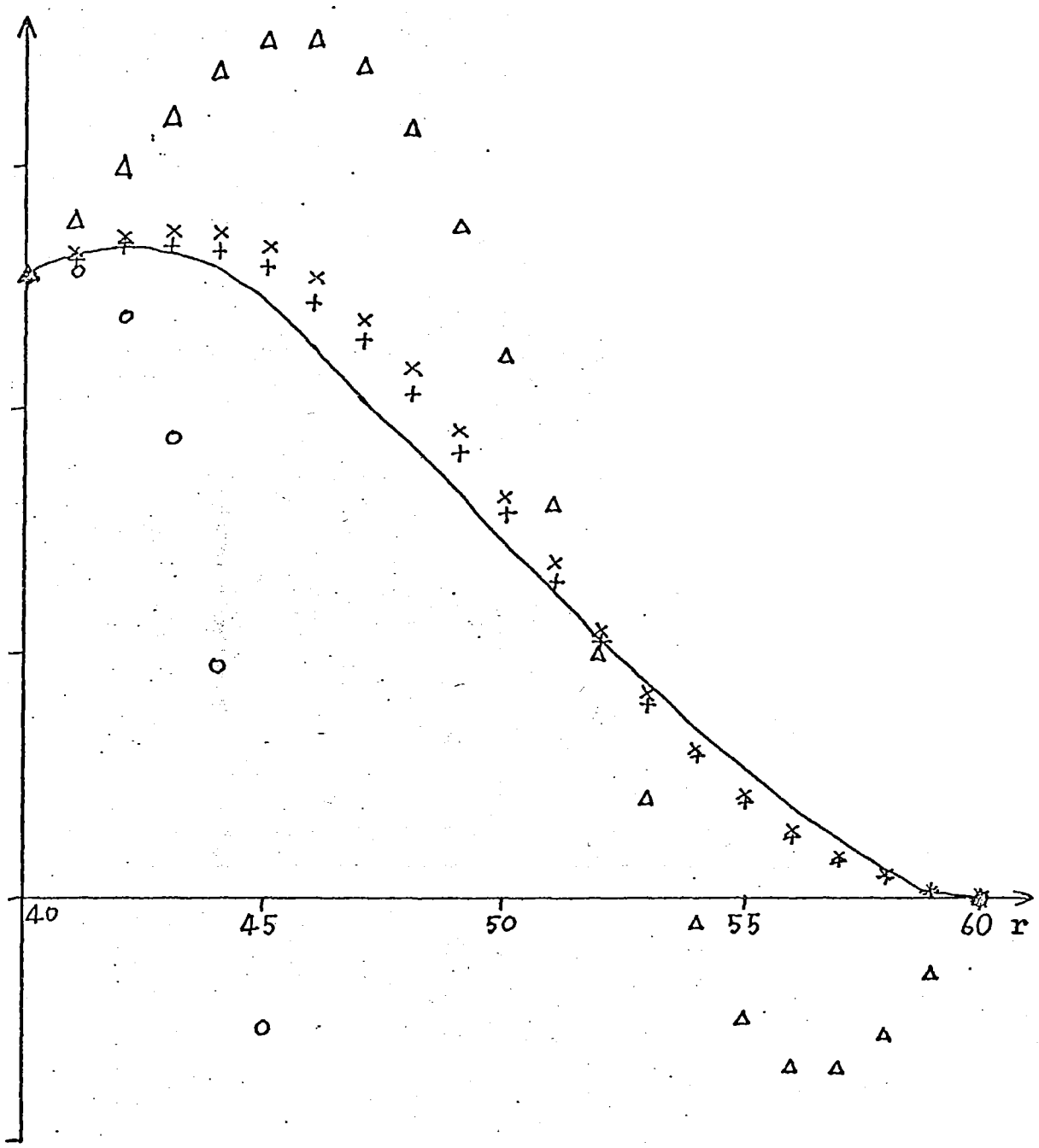


Fig.5.4 b estimated outside Γ; varying N_z with N=10.

——— actual b,

o $N_z = 0$,

Δ $N_z = 1$,

x $N_z = 2$,

+ $N_z = 4$.

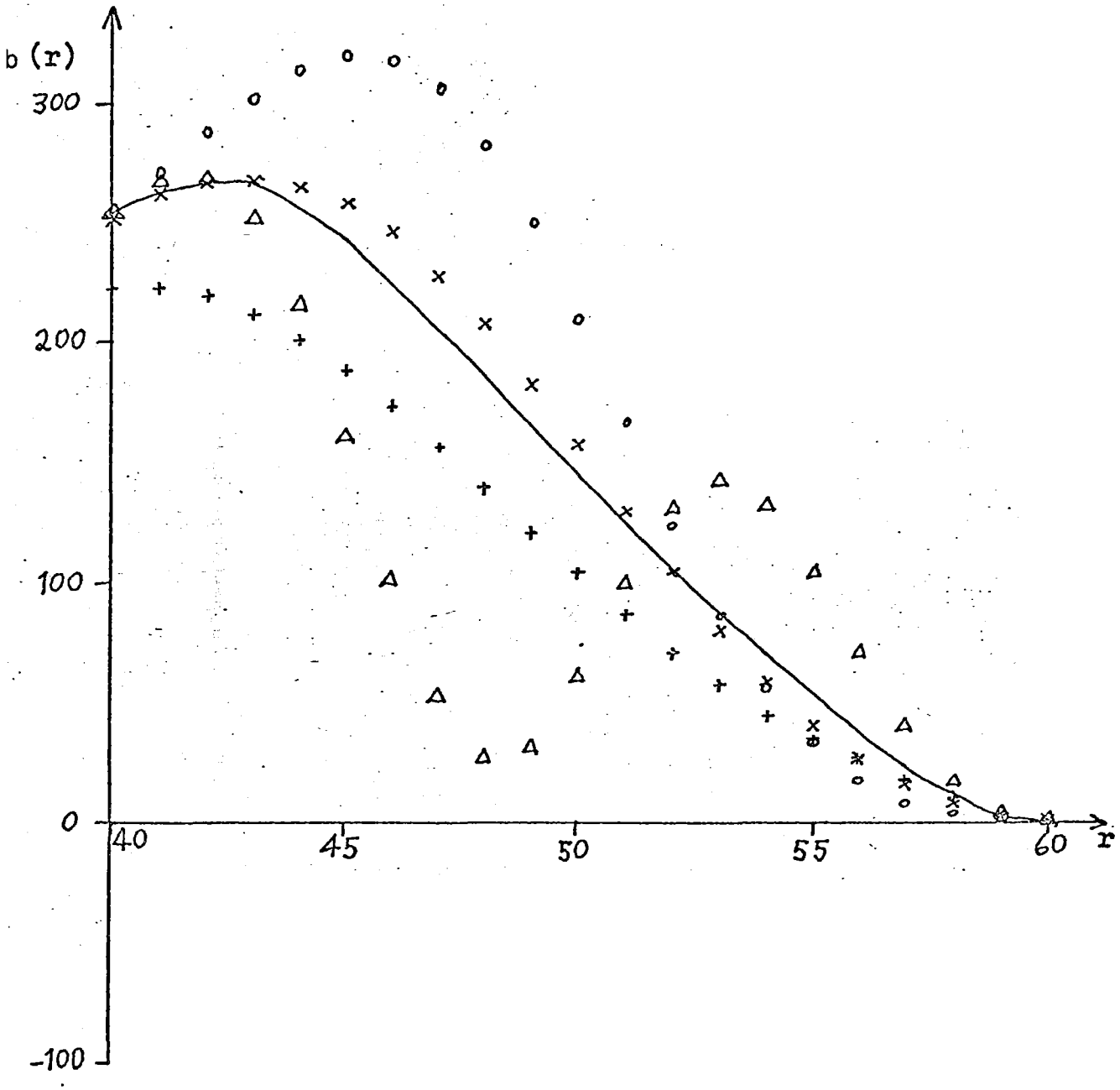


Fig. 5.5(a). b estimated outside Γ; varying N with $N_z = 4$.

——— actual b,

\+ N = 5,

x N = 10,

o N = 15,

Δ N = 20.

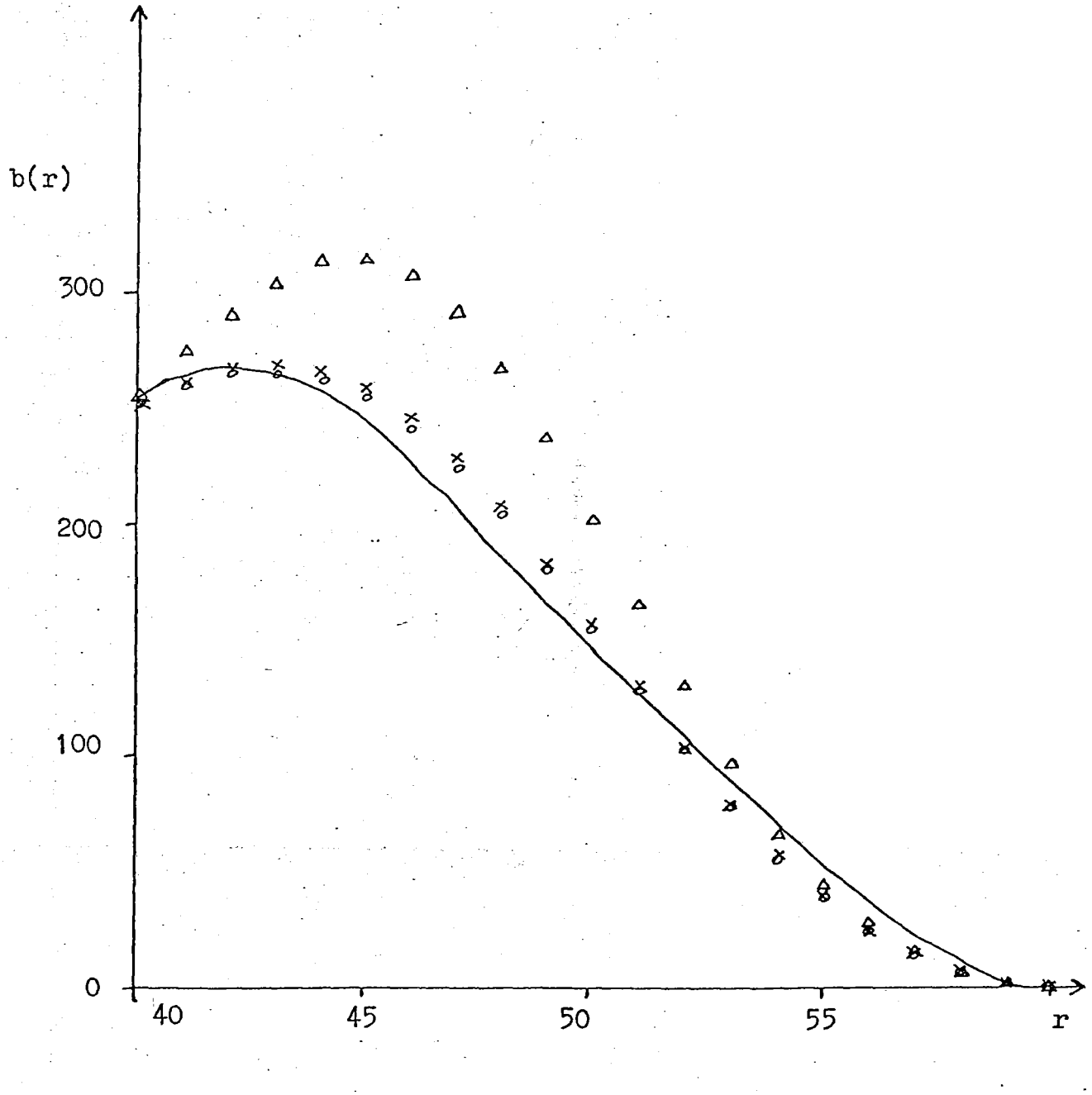


Fig.5.5(b). b estimated outside Γ; varying N with N_z = 4.

_____ actual b,

o N = 9,

x N = 10,

Δ N = 11.

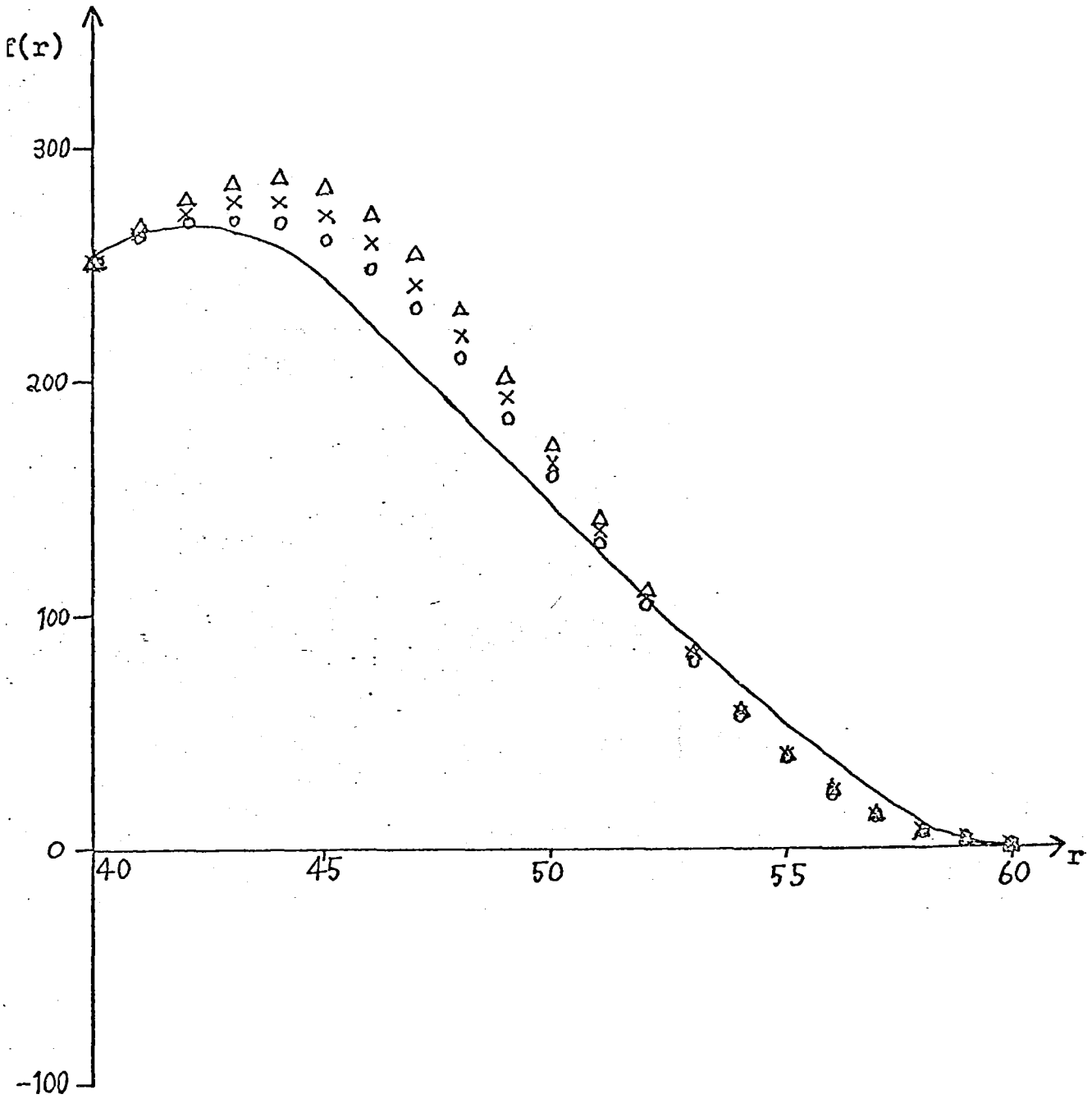


Fig. 5.6. f estimated outside Γ for varying noise levels on f in Γ, with $N_z = 4$ and $N = 10$.

_____ actual f,

o 1% noise,

x 5% noise,

Δ 10% noise.

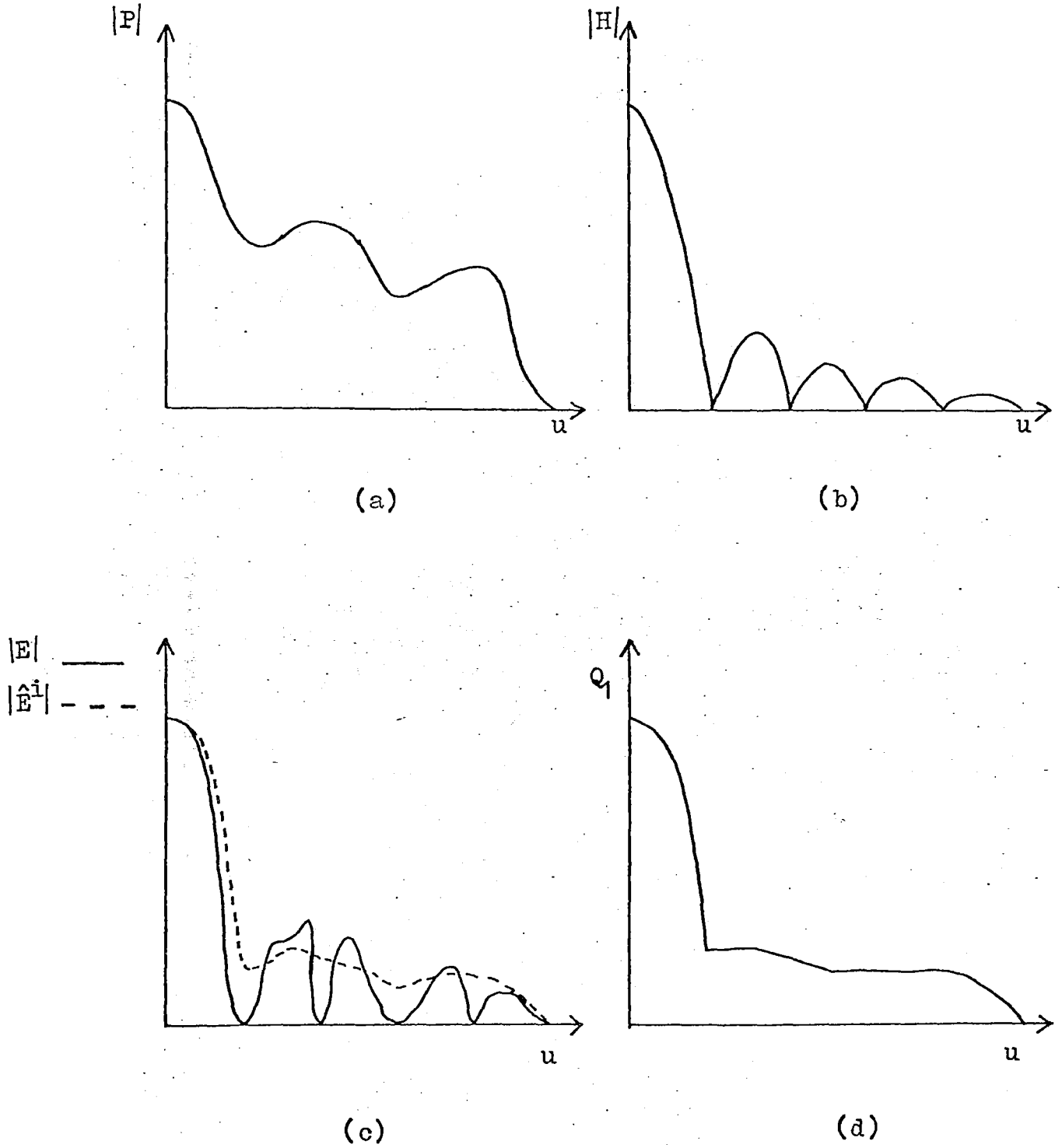
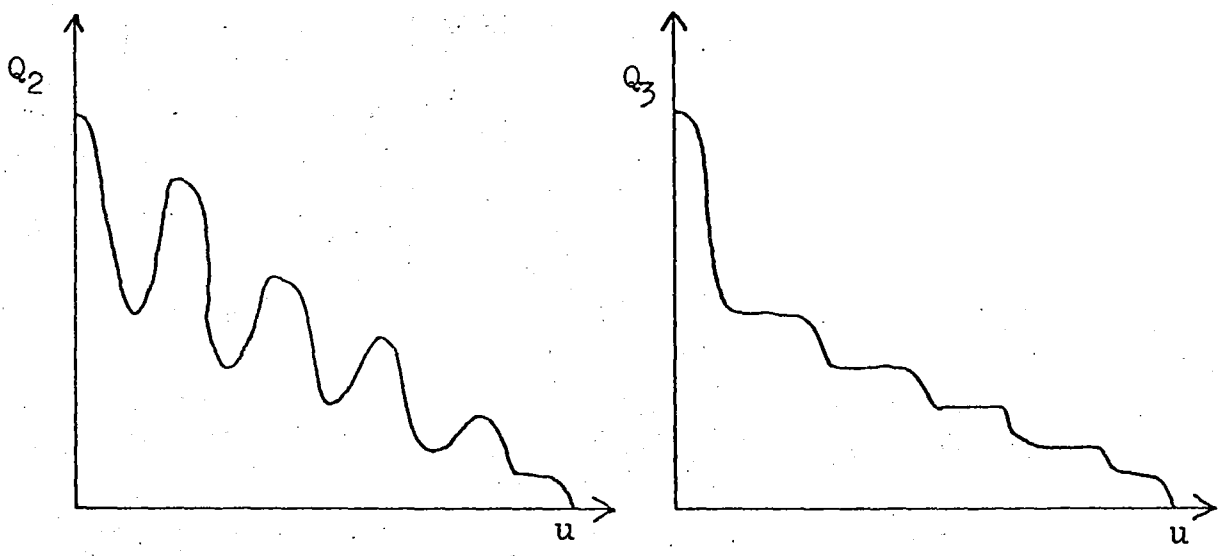


Fig.5.7 continued on next page.



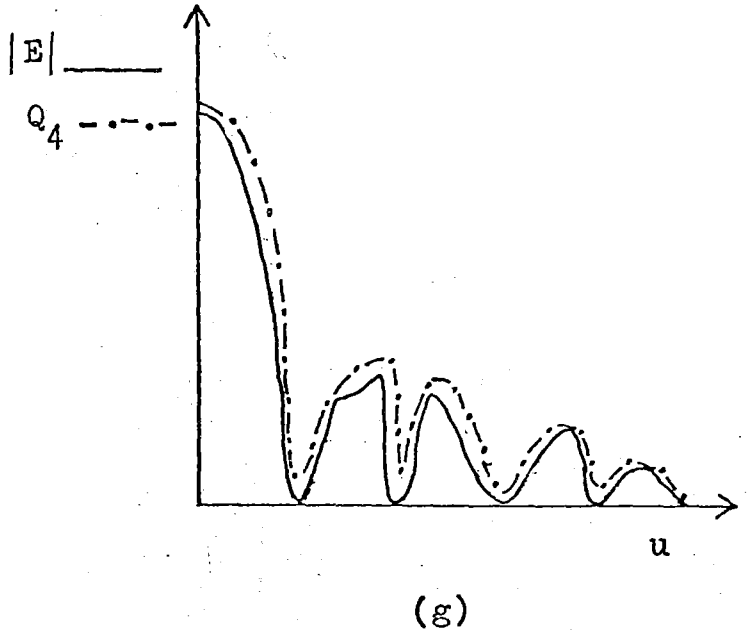
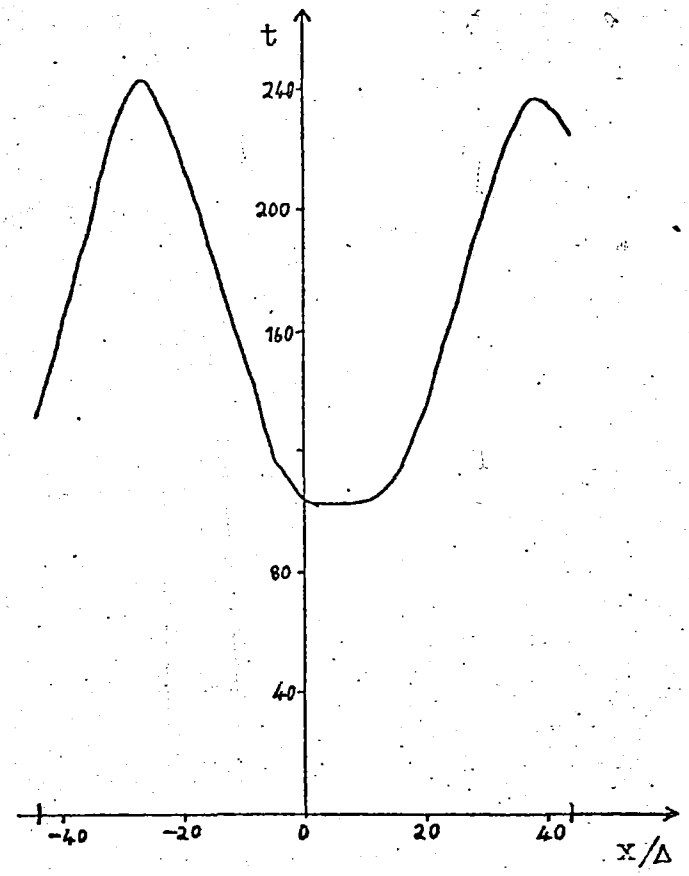
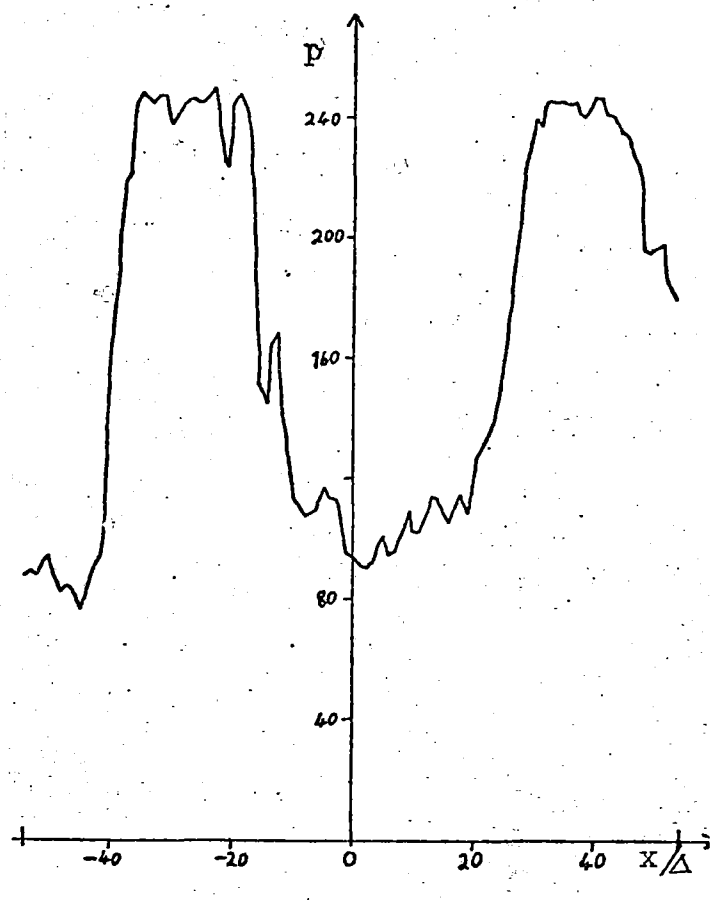


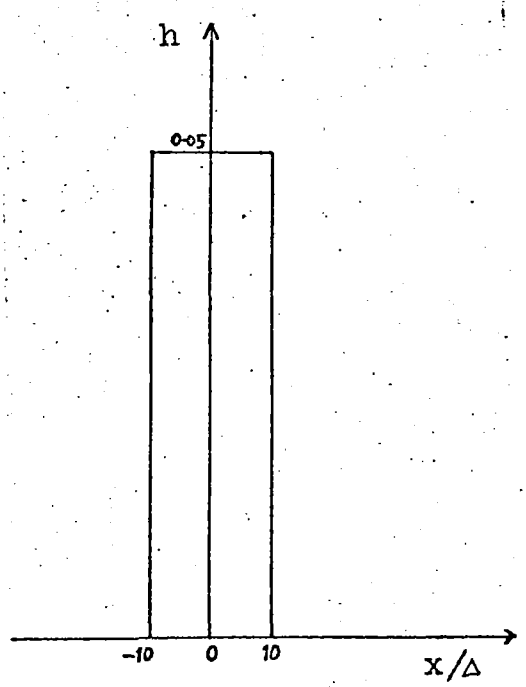
Fig.5.7. (a)|P|, (b)|H|, (c) ___|E|, _ _ _ $|\hat{E}^i|$, (d) Q_1, (e) Q_2, (f) Q_3, (g) ___|E|, -.-.- Q_4.



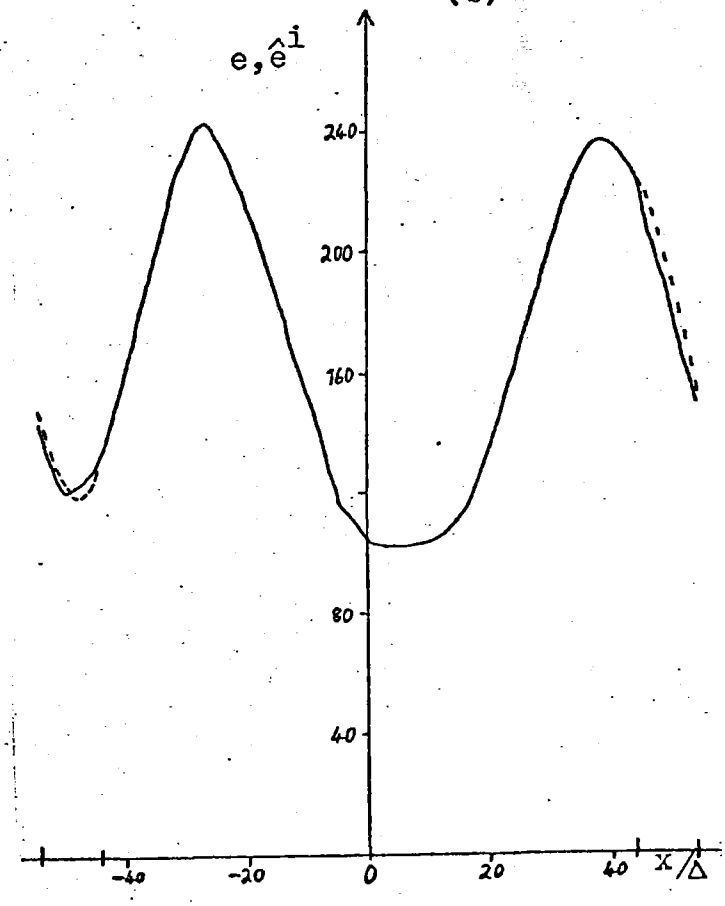
(a)



(b)

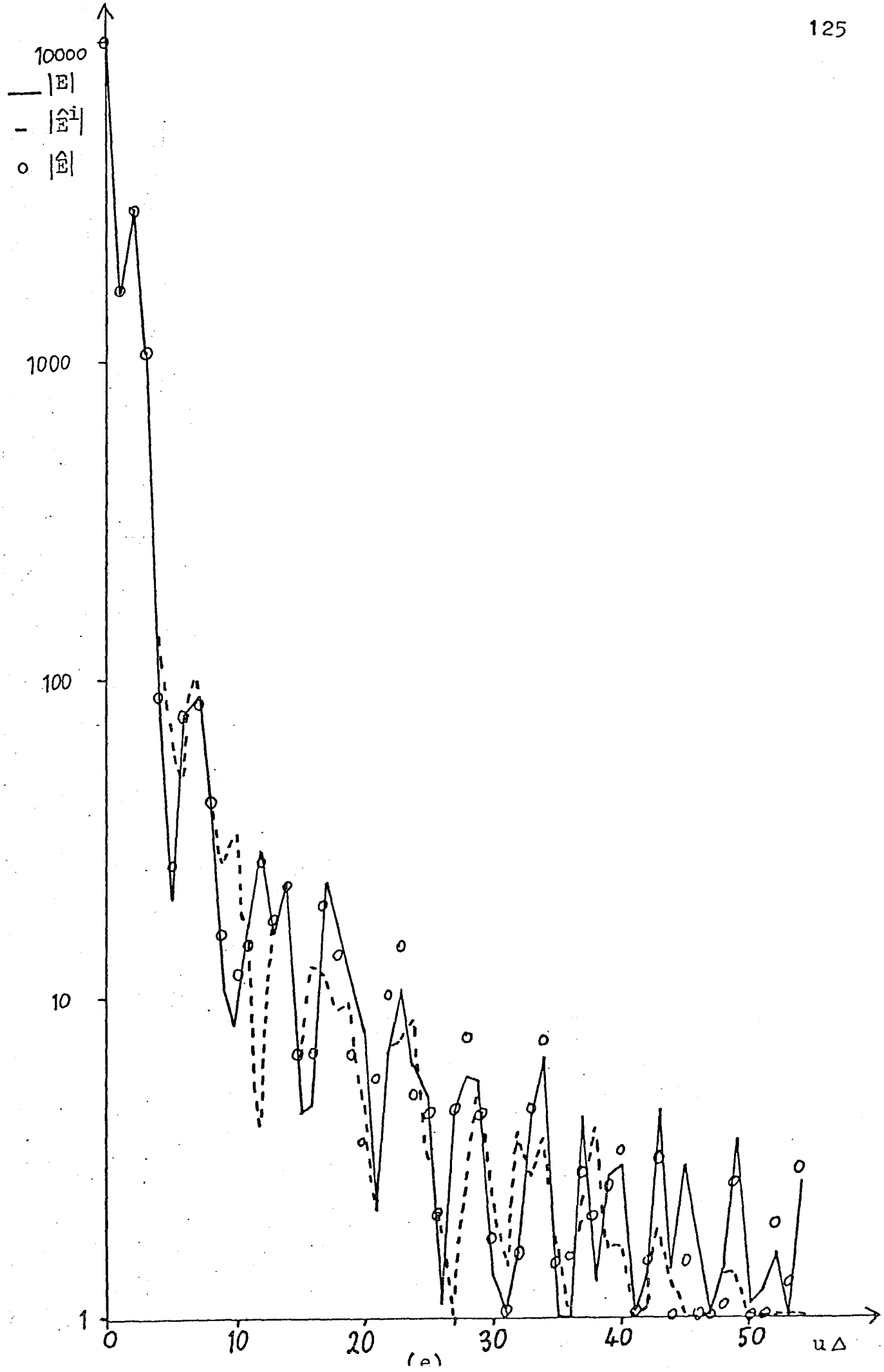


(c)

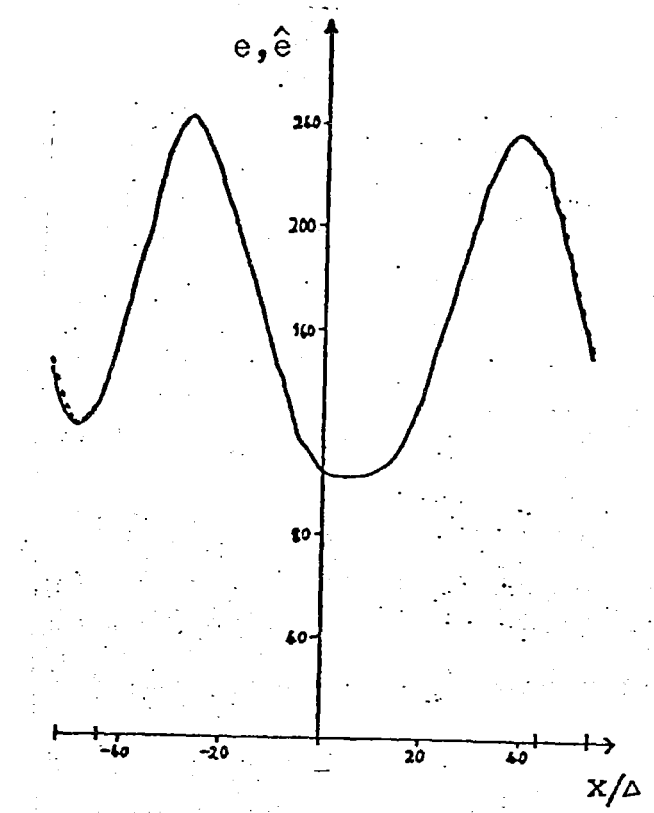


(d)

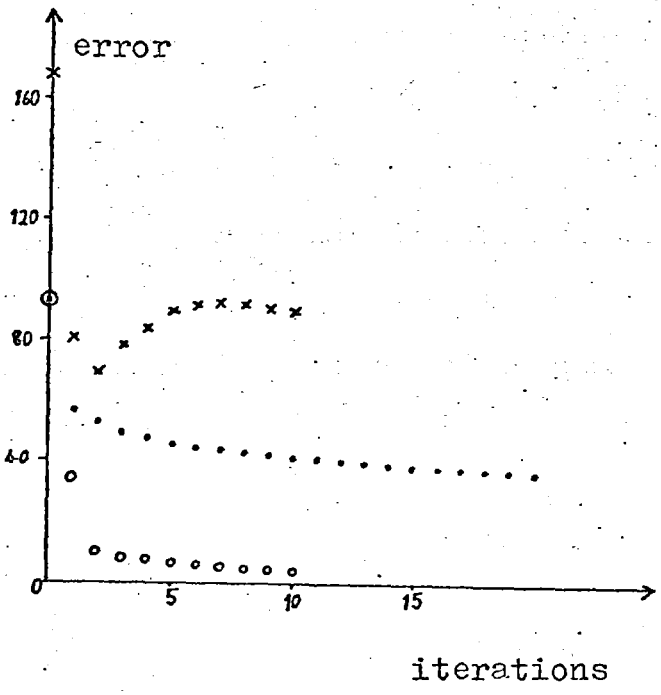
Fig.5.8 continued on next page.



(e)



(f)



(g)

Fig.5.8. (a) t, i.e. b within Γ, (b) p, (c) h, (d) e and $\hat{e}^i$; (e) $|E|$, $|E^i|$ and $|\hat{E}|$ after 10 iterations plotted on a logarithmic scale; (f) e and $\hat{e}$ after 10 iterations; (g) plot of error between $\hat{e}$ and e against the number of iterations for: . the noise free case, o the case when $E_{u_B} = |E|$, x in the presence of 3% noise.

CHAPTER SIX

NONRECURSIVE IMAGE RESTORATION USING A FINITE FILTER ARRAY

6.1 INTRODUCTION

It is pointed out in 1.4.2.2 that nonrecursive filtering can be accomplished directly by convolution with a finite filter array in image space. In this chapter, criteria used in deriving such filter arrays are reviewed and a new criterion is suggested which leads to arrays of shorter lengths. The effect of varying the filter array length, for a given degradation, is studied. The effect of different degradations on fixed filters are assessed. Consideration is given to the amplification of noise by the filter, and both 1-D and 2-D filters are dealt with. The analysis is illustrated with computational examples.

6.2 REVIEW OF CRITERIA FOR DERIVING A FINITE FILTER ARRAY

In 1.2 the image restoration problem is stated as follows: given f, obtain an estimate $\hat{p}$ of p. For simplicity, the problem is considered here in 1-D, except in 6.7 where it is shown how the theory may be simply extended to 2-D. In chapters 4 and 5, techniques are developed for overcoming the problems discussed in 2.3, which are associated with class $\mathscr{G}$ degraded images. The technique known as weighted superposition of images (Saleh 1974), or

discrete convolution (Arguello et al 1972, Frieden 1974), also overcomes the above problems when $f \in \mathcal{G}$, provided that the restoration requirements are relaxed, i.e. the restoration problem now becomes: given $f \in \mathcal{G}$, obtain an estimate $\hat{p}$ of p for $|x| \leqslant (A-L_1')/2$, where L_1' is the extent of a filter array which is to be convolved with f to provide $\hat{p}$.

Assume that sampled values of f are available, i.e. the set $\{f(x_m), m=-M,\ldots,M\}$ is given, where $x_m = m\Delta$, $A = 2M\Delta$ and Δ is the sampling distance. Let f be convolved with a discrete filter array $\underset{\sim}{W} = \{w_n, n=-N,\ldots,N\}$, where the array elements are again separated by Δ. Thus

$$\hat{p}(x) = \sum_{n=-N}^{N} w_n f(x - x_n), \quad |x| \leqslant (M-N)\Delta. \tag{6.1}$$

Taking (6.1), (2.1), (2.3) and (2.4) together gives

$$\hat{p} = b \otimes s + c', \quad |x| \leqslant (M-N)\Delta, \tag{6.2}$$

where

$$s = s(x) = \sum_{n=-N}^{N} w_n h(x - x_n) \tag{6.3}$$

and

$$c' = c'(x) = \sum_{n=-N}^{N} w_n c(x - x_n). \tag{6.4}$$

After convolution with $\underset{\sim}{W}$ the resultant psf and contamination are s and c' respectively. The problem now is to choose $\underset{\sim}{W}$ so as to optimise s in some way without unduly increasing the variance of c'.

In considering this problem, some authors (Frieden 1974, Saleh 1974) use a continuous psf and have thus tried to optimise a continuous s. Here, however, let $\underset{\sim}{W}$ be convolved with $h_s = h_s(x)$, where the latter is called the "sampled psf":

$$h_s(x) = \sum_{j=-J}^{J} h_j \, \delta(x - j\Delta), \tag{6.5}$$

where $h_j = h(j\Delta)$. This gives a sampled s

$$s(x) = \sum_{j=-(N+J)}^{N+J} s_j \, \delta(x - j\Delta), \tag{6.6}$$

where

$$s_j = \sum_{n=-N}^{N} w_n \, h_{j-n}. \tag{6.7}$$

For convenience let $h_j = 0$ for $|j| > J$. It is proposed to optimise the sampled s in (6.6) and then obtain a continuous s by linearly interpolating between the samples.

Various criteria (Arguello et al 1972, Nathan 1971, Honda et al 1974, Riemer et al 1973, Frieden 1974 and Saleh 1974) have been proposed for deriving $\underset{\sim}{W}$. Most of these (Honda et al 1974, Riemer et al 1973, Frieden 1974 and Saleh 1974) attempt to optimise s in some way. Clearly, it is desirable to find that filter array $\underset{\sim}{W}_I$, which is called "ideal", that forces s(x) to be a delta function, i.e.

$$s(x) = h(x) \otimes \underset{\sim}{W}_I = \delta(x). \tag{6.8}$$

However, (6.8) can never be satisfied exactly when N is finite, because S_{N+J} cannot be zero. Arguello et al (1972)

obtain $\underset{\sim}{W}$ by truncating, with a rectangular or Hamming window, the inverse FT of $\overline{H}$. A similar approach is taken by Nathan (1971). Honda et al (1974) obtain $\underset{\sim}{W}$ by minimising, up to some cutoff frequency, the mean square difference between the optical transfer functions of $\hat{p}$ and p. Frieden (1974) considers s to be symmetrical, forces it to have its first zero from the origin at a prescribed position, and finds the $\underset{\sim}{W}$ which minimises the maximum sidelobe. Following on from Smith (1966), Riemer and McGillem (1973) minimise the radius of gyration of s:

$$\int_{-\infty}^{\infty} q(x)\ s^2(x)\ dx \tag{6.9}$$

under the constraint

$$\int_{-\infty}^{\infty} s^2(x)\ dx = 1, \tag{6.10}$$

where $q = q(x)$ is a weighting function. Saleh (1974), setting $q(x) = x^2$, minimises

$$\int_{-\infty}^{\infty} x^2\ s^2(x)\ dx \tag{6.11}$$

under the condition that

$$\int_{-\infty}^{\infty} s(x)\ dx = 1. \tag{6.12}$$

Riemer and McGillem (1973) adopt a Fourier space approach to minimising (6.9) under the constraint (6.10). Their solution for $\underset{\sim}{W}$, though general, is considerably more

difficult to obtain than that of Saleh (1974), which simply involves a matrix inversion.

Because of the general possibility that s may be infinite for some values of x and yet may still satisfy (6.12), it seems that (6.10) is a more appealing constraint. In early work for this thesis, matrix techniques for minimising (6.11) under the constraint (6.10) were experimented with for various psf, and the results obtained were very similar to those obtained from Saleh's method. However, it was found that Saleh's method is simpler to use and tends to be more stable. It is the particular form of (6.11) that ensures the stability of s.

6.3 AN IMPROVED CRITERION

In discussing the choice of q, Riemer and McGillem (1973) suggest that q must increase rapidly away from the origin. This is because it is desirable to concentrate s near the origin. However, I feel that this places undue emphasis on values of s which may be small and distant from the origin. Such values could better be treated as additive noise. Therefore, s is now considered to be sampled and it is proposed to maximise s_0 under constraints (6.10), which now becomes

$$\sum_{j=-(N+J)}^{N+J} s_j^2 = 1. \qquad (6.13)$$

Let

$$G = s_0 - \frac{\mu}{2}\left(\sum_{j=-(N+J)}^{N+J} s_j^2 - 1 \right) \qquad (6.14)$$

where μ is a Lagrangian multiplier. G will be a maximum when

$$\frac{\partial G}{\partial w_k} = 0, \qquad k=-N,..,N \tag{6.15}$$

so that (6.7) and (6.14) in (6.15) give

$$\mu \sum_{n=-N}^{N} \sum_{j=-(N+J)}^{N+J} h_{j-n}\, h_{j-k}\, w_n = h_{-k}, \qquad k=-N,..,N. \tag{6.16}$$

Let $\underset{\sim}{R}$ be the matrix whose elements are given by

$$r_{k,n} = \sum_{j=-(N+J)}^{N+J} h_{j-n}\, h_{j-k}, \qquad n,k=1,..,N \tag{6.17}$$

and $\underset{\sim}{V}$ be the matrix whose elements are given by

$$v_k = h_{-k}, \qquad k=-N,..,N. \tag{6.18}$$

Then (6.16) becomes

$$\mu \underset{\sim}{R}\, \underset{\sim}{W} = \underset{\sim}{V}, \tag{6.19}$$

so that

$$\underset{\sim}{W} = \frac{1}{\mu} \underset{\sim}{R}^{-1}\, \underset{\sim}{V}. \tag{6.20}$$

Solving (6.20) gives $\underset{\sim}{W}$, where μ is a normalisation constant.

As N increases, s_0 must be monotonically increasing. This is because (by setting $w_N = w_{-N} = 0$) each possible $\underset{\sim}{W}$ of length (N-2) is included in the possible $\underset{\sim}{W}$'s of length N. Thus as N becomes large, s_0 approaches unity, which is the required property of $\underset{\sim}{W}$.

It is interesting to note that maximising s_0 in (6.14)

is equivalent to minimising

$$\sum_{j=1}^{N+J} s_j^2. \qquad (6.21)$$

This is because s_0 is maximum when s_0^2 is maximum, and s_0^2 is maximum when (6.21) is a minimum because of the normalisation in (6.13). This is a useful property of the improved criterion discussed in this section.

In deriving the $\underset{\sim}{W}$ given by (6.20), the central sample s_0 of the resultant psf has been maximised. It is not necessarily true that s_0 is the best sample of s to maximise, unless h is symmetrical about h_0. However, if h is unsymmetrical, it seems probable that an improved s may be obtained by maximising some sample s_l of s where $l \neq 0$. It is not difficult to adapt the foregoing theory to include this possibility. The solution is the same as (6.20) except that (6.18) becomes

$$v_k = h_{l-k}, \qquad k=-N,..,N. \qquad (6.22)$$

The best $\underset{\sim}{W}$ may be found by first evaluating, for each l, the corresponding $\underset{\sim}{W}$ given by (6.20) and then the s_l given by (6.7). The $\underset{\sim}{W}$ finally chosen is the one which gives the maximum s_l. Note that this procedure only involves a single inversion of $\underset{\sim}{R}$ in (6.20).

In the examples presented later in this chapter it is s_0 which is maximised because h is symmetrical.

6.4 NOISE EFFECTS

The presence of contamination or noise on f is inevitable. Convolving with a filter array usually amplifies any noise that is present. Therefore, the increase in s_0 obtained using a filter $\underset{\sim}{W}$ must be balanced with the noise amplification which the filter produces. Assume that the noise components n_j on different samples $f(x_j)$ of f have the same statistics but are uncorrelated. The value n'_j of the noise at the j'th sample of the filtered image is given by

$$n'_j = \sum_{k=-N}^{N} w_k \, n_{j-k}. \tag{6.23}$$

Let the expectation (i.e. mean) of the noise components be

$$E(n_j^2) = \sigma_n^2. \tag{6.24}$$

Then

$$\sigma_{n'}^2 = E(n_j'^2) = \sigma_n^2 \sum_{k=-N}^{N} w_k^2 = \sigma_n^2 \, C_n^2 \tag{6.25}$$

where C_n is the factor by which the noise level is multiplied. Also the correlation coefficient (Arguello *et al* 1972) between n'_j and n'_k is

$$\rho_{j,k} = \frac{E(n'_j \, n'_k)}{\sigma_n'^2} \tag{6.26}$$

$$= \frac{\sigma_n^2}{\sigma_{n'}^2} \sum_{l=-N}^{N} w_l \, w_{l+k-j} \, . \tag{6.27}$$

Thus, not only is the noise level multiplied by C_n but also the noise components of different samples of the image become correlated.

The effect of noise can be taken into account in deriving $\underset{\sim}{W}$ by adding

$$\lambda \sum_{k=-N}^{N} w_k^2 \tag{6.28}$$

to G in (6.14) and proceeding as before, where λ is a second Lagrangian multiplier. Riemer and McGillem (1973) use a similar term to constrain the noise amplification. A similar but more convenient approach is taken by Saleh (1974). He replaces G in (6.14) by G' where

$$G' = G\cos(\theta) + \left(\sum_{k=-N}^{N} w_k^2\right)\sin(\theta). \tag{6.29}$$

Noise amplification is reduced by increasing θ from 0 to $\pi/2$. This is the approach adopted in the examples in 6.6. It leads to the derivation of the best $\underset{\sim}{W}$ for a given value of the integer N (introduced in 6.2), consistent with the noise level present on f.

6.5 SAMPLING CONSIDERATIONS

If the sampling rate $1/\Delta$ is too slow, quantisation errors can lead to a poor restoration. Conventionally, f is assumed to have been bandlimited by the imaging system and is sampled at the corresponding Nyquist rate (Frieden 1974, Saleh 1974). However, this rate can be considerably faster

than the rate ($1/\Delta_a$ say) required to display adequately the detail of interest in p. My experience is that a certain amount of a priori information is almost always available for use in the restoration procedure. This information includes an estimate not only of the psf, but also of $1/\Delta_a$, of the maximum noise amplification C_n^{max} which $\underset{\sim}{W}$ can be allowed to generate, and of the maximum acceptable noise level σ_n^{max} on $\hat{p}$. I feel that in the context of image restoration by discrete convolution the following criteria lead to a sampling rate more suitable than the Nyquist rate. Firstly, the sampling rate should be at least $1/\Delta_a$. Secondly, it should be sufficient to ensure that the continuous function $h_c = h_c(x)$, called the "line segment psf", obtained by linearly interpolating between the samples $\{h_j, j=-J,..,J\}$ of h_s is a satisfactory approximation to h itself. By "satisfactory" it is meant that the noise level

$$\sigma_h = \frac{1}{\sqrt{2J+1}} \int_{-\infty}^{\infty} \left| h(x) - h_c(x) \right| dx \qquad (6.30)$$

introduced by the line segment approximation h_c to h will not, when convolved with $\underset{\sim}{W}$, give rise to a noise level in $\hat{p}$ which is greater than σ_n^{max}. In other words, it is required that

$$\sigma_h \, C_n^{max} < \sigma_n^{max} \, . \qquad (6.31)$$

The effect of the above sampling criteria is now illustrated. In doing this it is sufficient to consider the ideal filter array $\underset{\sim}{W}_I$. This is because the difference

between $\underset{\sim}{W}$ and $\underset{\sim}{W}_I$ gives rise to what is in effect a background noise level in $\hat{p}$. To develop the argument it is necessary to define the triangular psf $s_1 = s_1(x)$ shown in Fig. 6.1(a).

When deriving $\underset{\sim}{W}$ in 6.3, h (see Fig. 6.1(b)) is taken to consist of a series of δ functions lying on sampling points (see (6.5) and Fig. 6.1(c)). After convolving with the optimum array $\underset{\sim}{W}$ of length 4Δ (Fig. 6.1(d)) an improved psf s (Fig. 6.1(e)) is obtained still in the form of a series of δ functions. This is treated as the sum of the ideal central δ function in (6.8) and noise. $\underset{\sim}{W}$ has been derived using samples of h obtained with h centred on a sampling point, as in Fig. 6.1(b). In an actual image, h need not be centred on a sampling point. Hence, it is necessary to know the effect on s in Fig. 6.1(e) of shifting h by a fraction of Δ. First replace h by the line segment psf h_c (Fig. 6.1(f)). Now, if $h_c(x)$ is replaced by $h_c(x-\alpha\Delta)$, it has the same sampled values as

$$(1 - \alpha)\ h_c(x) \quad + \quad \alpha\ h_c(x-\Delta) \tag{6.32}$$

which is shown in Fig. 6.1(h). Hence, $h_c(x-\alpha\Delta)$ is equivalent to the sum of two line segment psf centred on adjacent sampling points, as in Fig. 6.1(h) where $\alpha = 1/3$. Consequently, when $h_c(x-\alpha\Delta)$ is convolved with $\underset{\sim}{W}_I$; the result is two δ functions at adjacent sampling points whose magnitudes are inversely proportional to their distance from the true value on which h_c was centred. If h_c is replaced by s_1 then, the magnitudes of these δ functions are the

samples $s_1(-\alpha\Delta)$ and $s_1((1-\alpha)\Delta)$ of $s_1(x-\alpha\Delta)$.

This result is very satisfactory. It means that the ideal restored image at a given sampling rate is the result of sampling $p \otimes s_1$. Because the width of s_1 is only 2Δ, the ideal restored image is no worse than the result of directly sampling p. If the sampling rate is high enough to preserve the detail of interest in p then the line segment psf will be ideally suited to the restoration procedure in 6.3. Note that in this chapter, as in Arguello et al (1972), no assumption has been made about the bandlimitedness of p or h. The sampling concepts briefly introduced in this section are generalised and considered in detail in chapter 7.

6.6 ONE-DIMENSIONAL RESULTS AND DISCUSSION

Results obtained using the method described in 6.3 are now compared with those obtained using two other methods, namely (a) the method of Arguello et al (1972) using rectangular windows and normalising according to (6.10), and (b) the method of Saleh (1974) normalising according to (6.12). In order to make fair comparisons, final results are normalised according to (6.10).

In Fig. 6.2 are shown two particular line segment psf's. Fig. 6.3 shows the corresponding plots of s_0 against N for the three methods being considered. Fig. 6.4 shows the $\underset{\sim}{W}$ given by each method for the psf in Fig. 6.2(a) when N = 13. Fig. 6.5 shows the s corresponding to each $\underset{\sim}{W}$ in Fig. 6.4. Fig. 6.6 shows a plot of C_n against N for each method for the psf in Fig. 6.2(a).

The results shown in Figs 6.2-6.6 vary considerably as the psf is varied. However, they follow the general trend shown by these examples. The value of s_0 in Fig. 6.3 obtained by the method in 6.3 tends to level off as N increases at a value of N which varies considerably with the psf.

The results obtained by the method in 6.3 and that of Saleh are generally superior to those obtained by the method of Arguello *et al*. This is because Arguello *et al* do not use an optimisation procedure in deriving $\underset{\sim}{W}$. According to the criterion adopted in (6.14) for maximising s_0, the results using the method in 6.3 are superior to those of Saleh. For the reasons discussed in 6.3, I consider the criterion used in (6.14) to be preferable to that used by Saleh. In order to obtain a particular value of s_0 in Fig. 5, the method in 6.3 allows the shortest $\underset{\sim}{W}$ to be used. As N increases, for a particular value of s_0, C_n can be made to decrease by using the methods discussed in 6.4. The tradeoff between s_0 and C_n as θ varies in (6.29) is illustrated in Fig. 6.7. In Fig. 6.7(a) is shown the variation of C_n as a function of θ, for the psf in Fig. 6.2(a) and for N = 10. The corresponding plot of s_0 against θ is shown in Fig. 6.7(b).

It is appropriate here to discuss further the normalisation procedure adopted in deriving the filter array $\underset{\sim}{W}$ in 6.3. As derived in 6.3, $\underset{\sim}{W}$ maximises s_0, where s is normalised as in (6.13). If s had been normalised so that

$$\sum_{j=-N+J}^{N+J} S_j = 1, \qquad (6.33)$$

there would have been no stable solution for $\underset{\sim}{W}$. Consequently it was necessary to use the normalisation constraint (6.13) in deriving $\underset{\sim}{W}$. However, when using filter arrays in practice it is often desirable that convolution with $\underset{\sim}{W}$ should leave unchanged a function which is uniformly constant. This requires that s is normalised as in (6.33). If this is necessary, the $\underset{\sim}{W}$ derived as in 6.3 may be multiplied by a constant such that s is normalised according to (6.33). Once this is done, s_0 need no longer be monotonically increasing with N. However, the general trend is still that s_0 increases to 1 as N increases.

In Fig. 6.8 the performance of the three methods is compared on a line of an image blurred in the computer. Fig. 6.8(a) is a plot of the intensity levels of a line of an image of extent 136Δ. Fig. 6.8(b) shows the class $\mathcal{G}$ image (of extent 128Δ) resulting from the convolution of the function shown in Fig. 6.8(a) with the psf in Fig. 6.8(b) in the absence of noise. The restorations of extent 100Δ for N = 14 obtained using the three methods are shown in Figs 6.8(c)-(e). Because the degraded image is of class $\mathcal{G}$, only the central portion is restored. The sum of the moduli of the differences between the samples of the restored image and the corresponding samples of the original image are 1.34, 1.84, and 3.25 for Fig. 6.8(c), Fig. 6.8(d) and Fig. 6.8(e) respectively. The images have been scaled to have the same mean and range as the original.

To give an indication of how $\underset{\sim}{W}$ and the s_j in (6.6) vary as N increases, $\underset{\sim}{W}$ and the s_j are tabulated in Tables

6.1 and 6.2 respectively for the psf shown in Fig. 6.2(a) and for N ranging from 1 to 10.

6.7 TWO-DIMENSIONAL RESULTS AND DISCUSSION

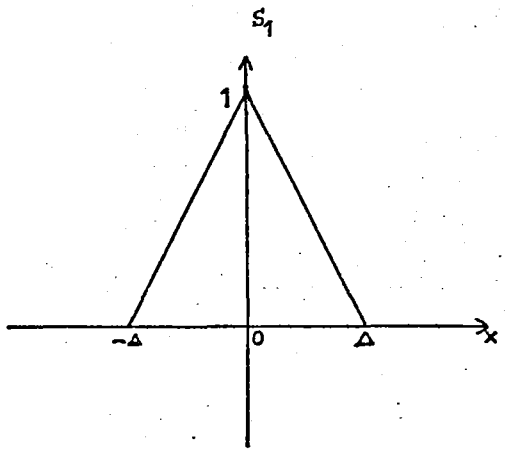
The theory and results presented so far in this chapter are for the case of a 1-D psf. As is discussed by Saleh (1974), it is straightforward to extend the theory in 6.3 to give an optimum two-dimensional filter array positioned on a rectangular grid. The dimension of the matrix corresponding to $\underset{\sim}{R}$ in (6.20) increases, but the general results remain the same. For example, in order to derive an optimum 9 x 9 element filter array it is necessary to invert an 81 x 81 matrix corresponding to $\underset{\sim}{R}$ in (6.20).

In Fig. 6.9(a) s_0 is plotted against N for the case of a $(2N + 1) \times (2N+1)$ filter array derived to optimise a psf which is a disc of radius 4Δ. In Fig. 6.9(b) the corresponding noise amplification factor C_n is plotted against N. These results indicate that the general properties of the 2-D filter arrays are similar to those of the 1-D filter arrays in 6.6.

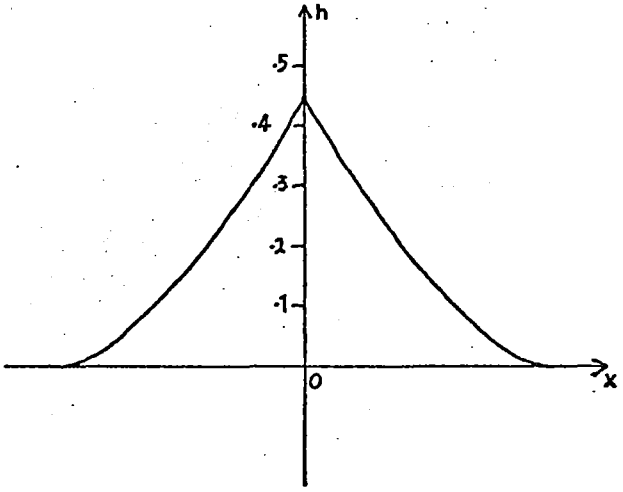
The 1-D nonrecursive filter arrays derived in 6.6 are suitable for use as real-time digital filters. 2-D filter arrays can also be used in this way, but their use is more complicated. As discussed by Huang et al (1971) a 2-D filter array can be decomposed into a number of 1-D filters which may be simpler to implement.

In general, image restoration results obtained using the edge extension method of chapter 4 are superior to those

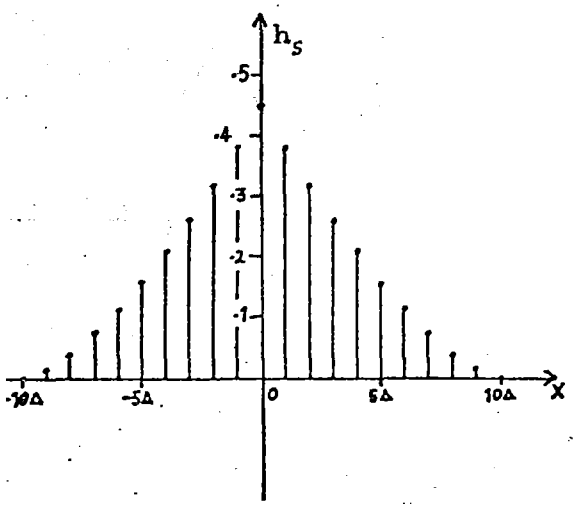
obtained using a nonrecursive filter array of a given length obtained by the method presented in 6.3. This is because the inverse psf $\bar{h}$ involved in the edge extension method is effectively of infinite extent because of the circular property of the FFT. However, Fig. 6.3 shows that as N increases, the difference between the results given by the two methods can become very small. Another disadvantage of the nonrecursive filtering technique introduced in this chapter is that the method of dealing with noise described in 6.4 is not as versatile as Wiener filtering. The main advantages of the method described in 6.3, as compared to the edge extension method, are its increased speed (if N is small enough) and its ease of implementation.



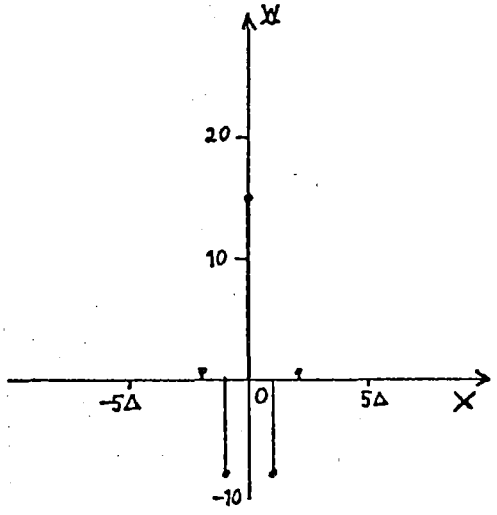
(a)



(b)

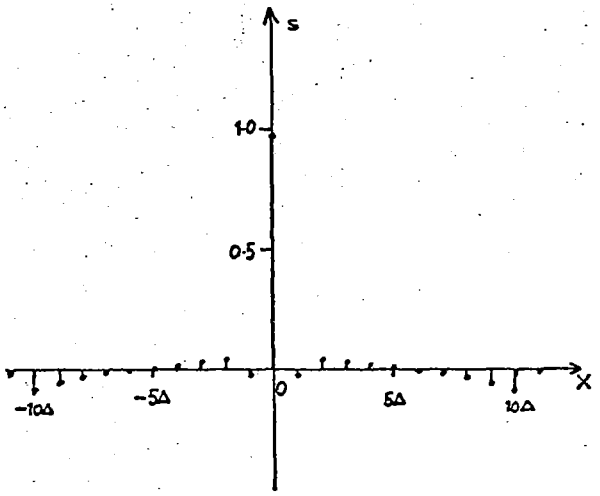


(c)

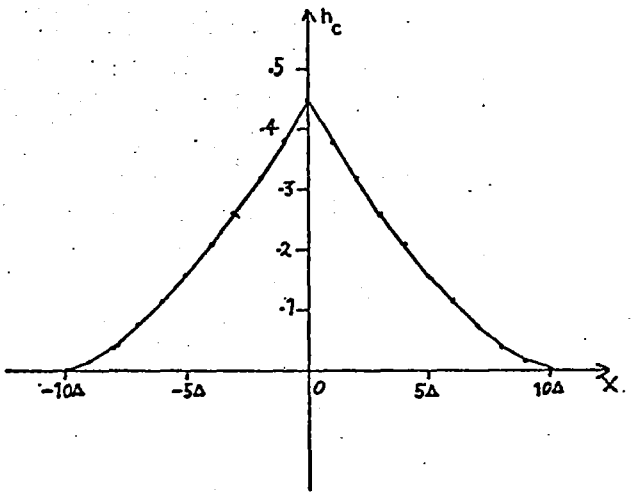


(d)

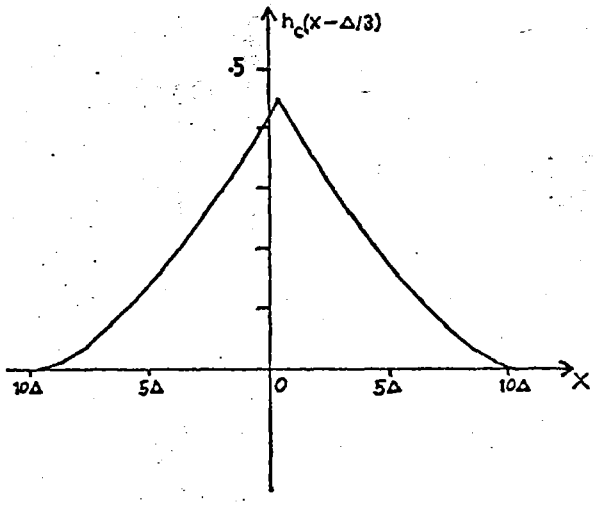
Fig. 6.1 continued on next page.



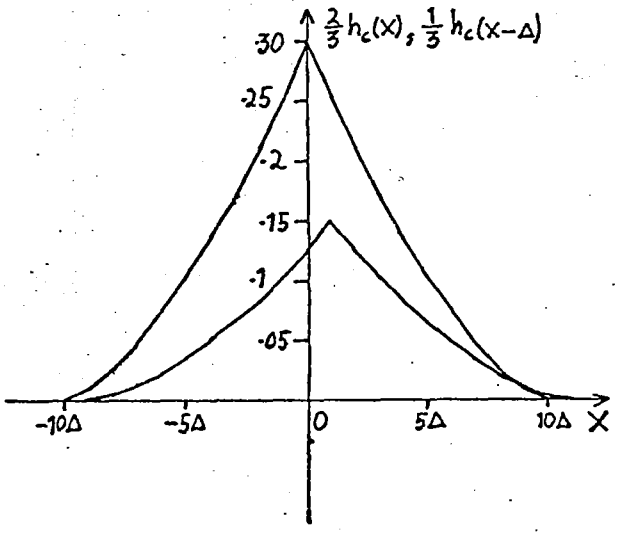
(e)



(f)

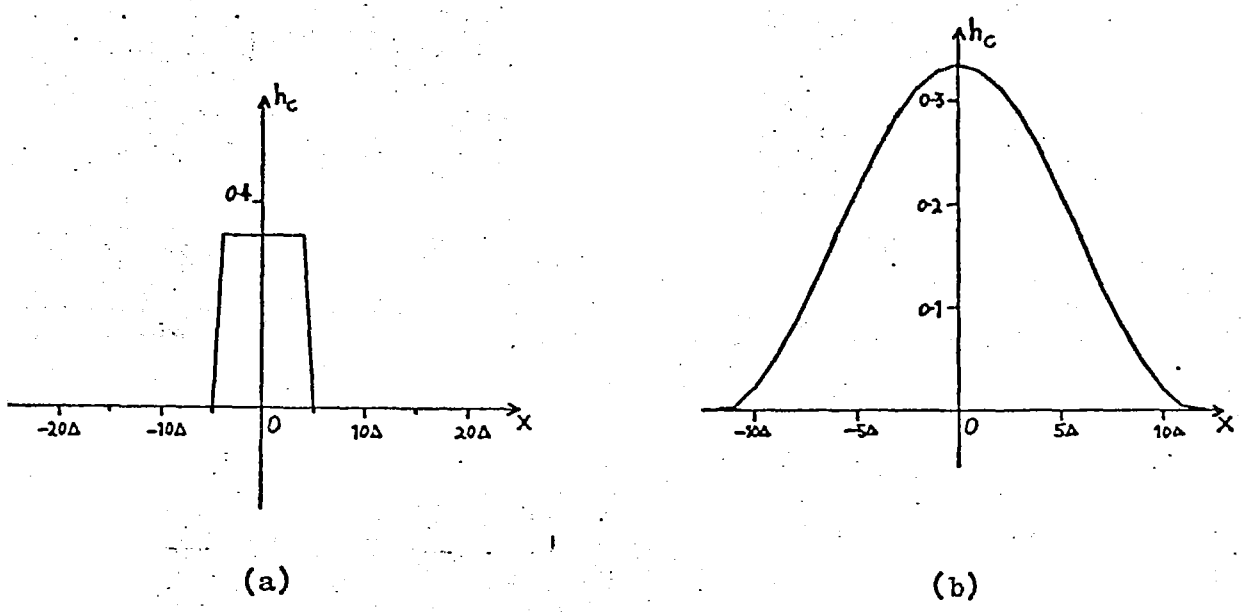


(g)



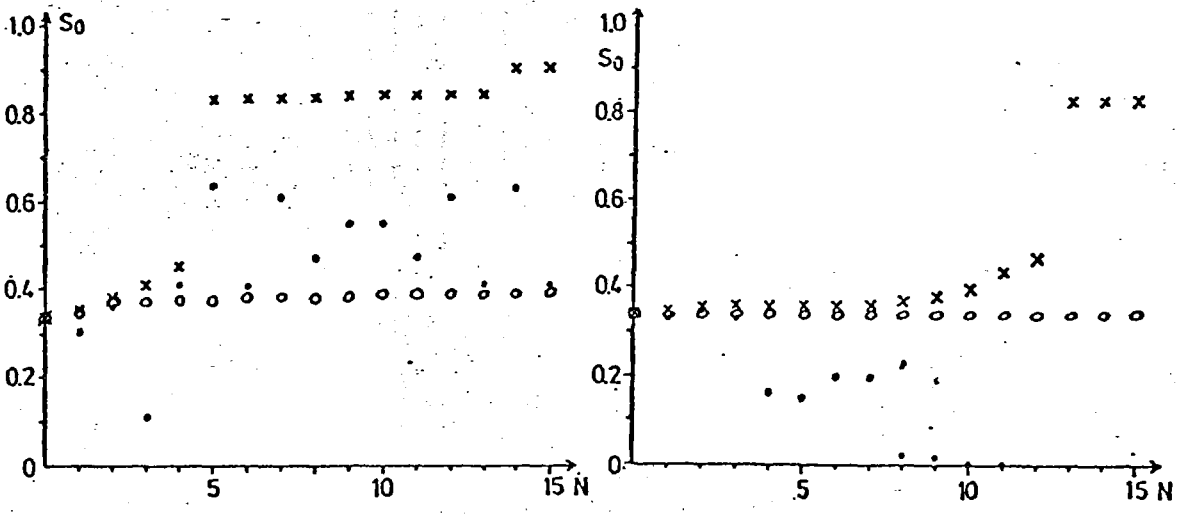
(h)

Fig. 6.1. (a) s_1, (b) h, (c) h_s, (d) $\underline{W}$, (e) s, (f) $h_c(x)$, (g) $h_c(x - \Delta/3)$, (h) 2/3 $h_c(x)$ and 1/3 $h_c(x - \Delta)$.



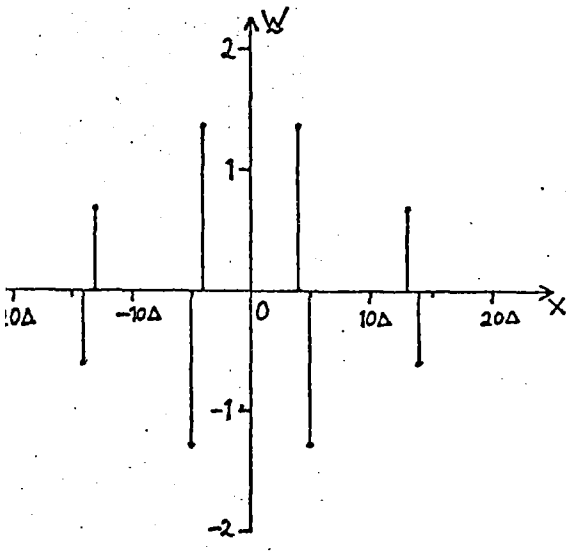
(a) (b)

Fig. 6.2. h_c obtained from (a) $h(x) = 1/3$, $|x| \leq 4\Delta$, (b) $h(x) = (1 + \cos(\pi x/(12\Delta))/6$, $|x| \leq 12\Delta$.

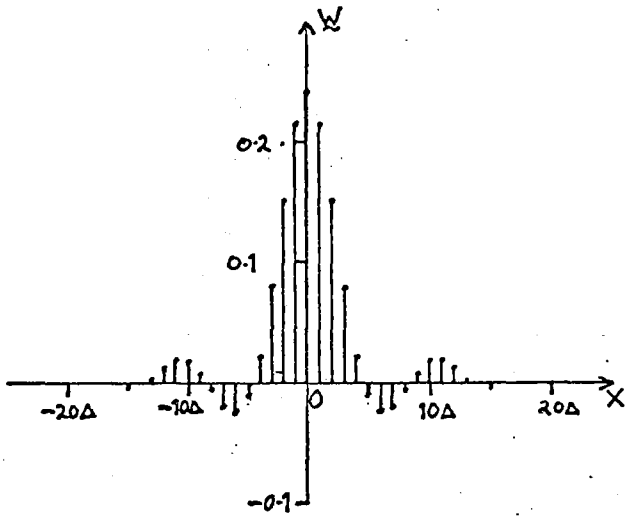


(a) (b)

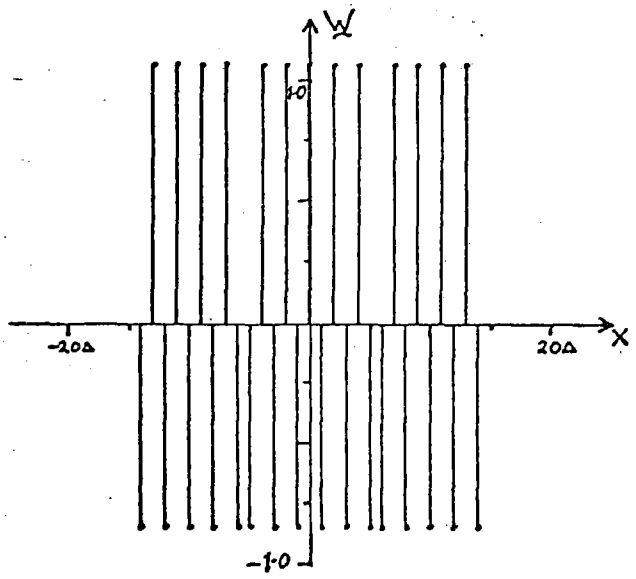
Fig. 6.3. (a) s_0 vs N for h_c in Fig. 6.2(a), (b) s_0 vs N for h_c in Fig. 6.2(b). Results are shown for our method (×), the method of Saleh (o), and the method of Arguello et al (•).



(a)

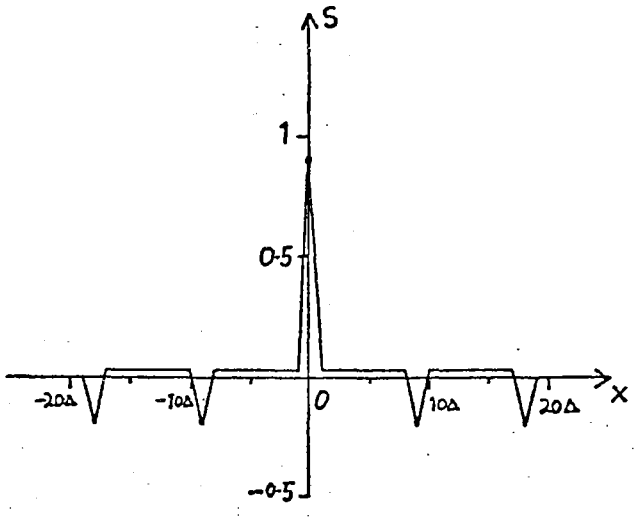


(b)

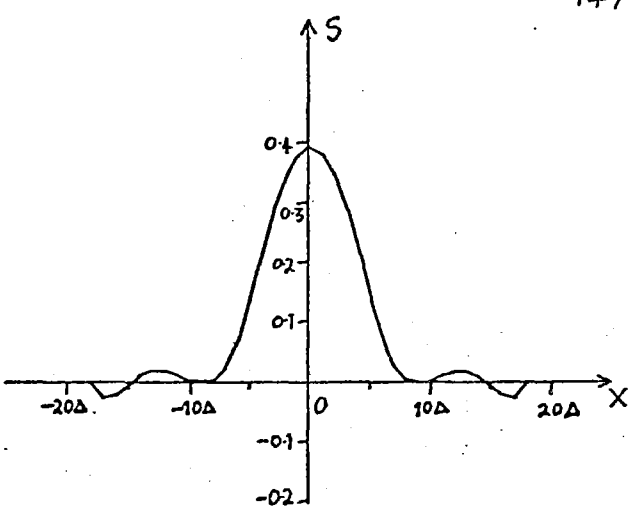


(c)

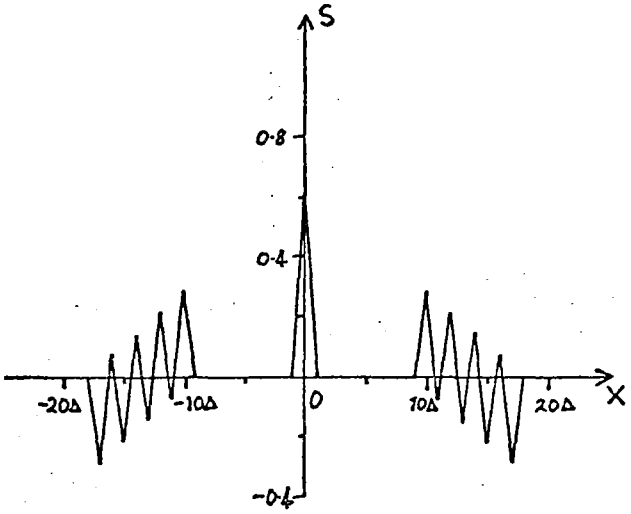
Fig. 6.4. $\underset{\sim}{W}$ corresponding to h_c in Fig. 6.1(a), for N=13, calculated by (a) our method, (b) method of Saleh, (c) method of Arguello et al.



(a)



(b)



(c)

Fig.6.5. (a,b,c) s corresponding to h_c in Fig.6.2(a) and respective values of $\underset{\sim}{W}$ in Fig. 6.4.

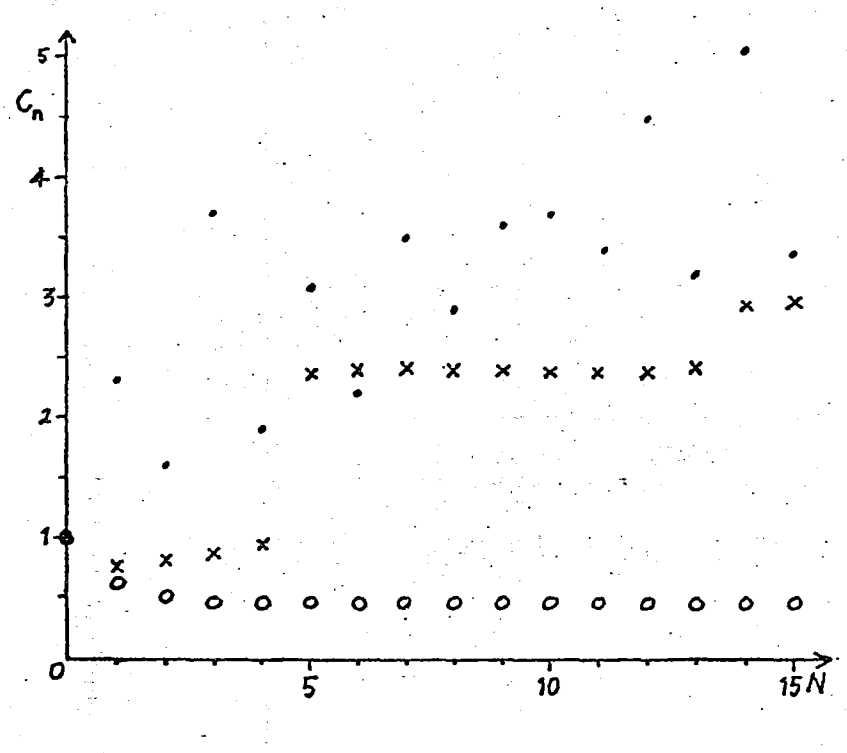
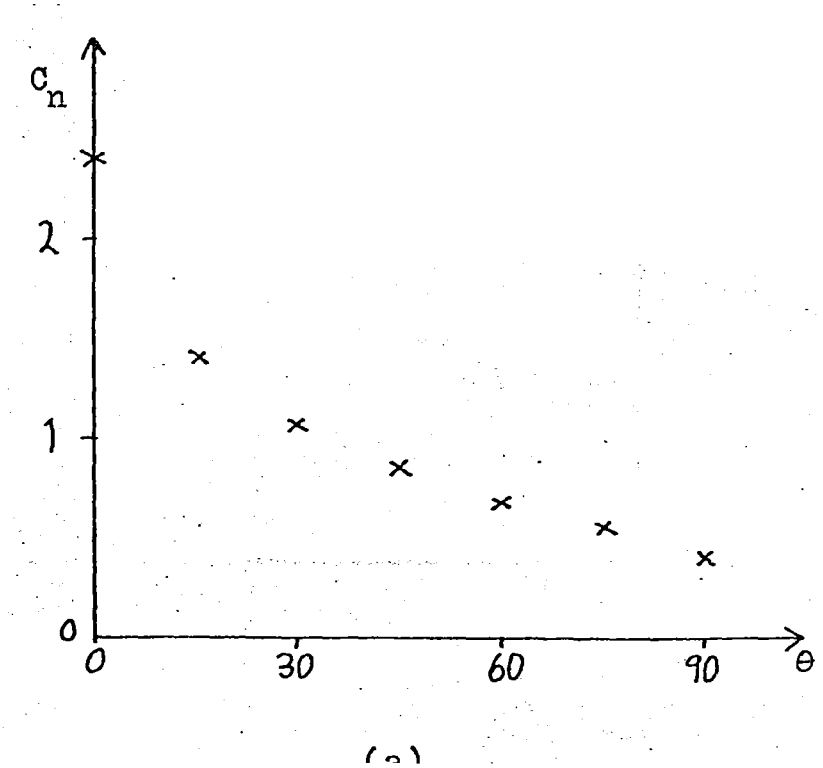
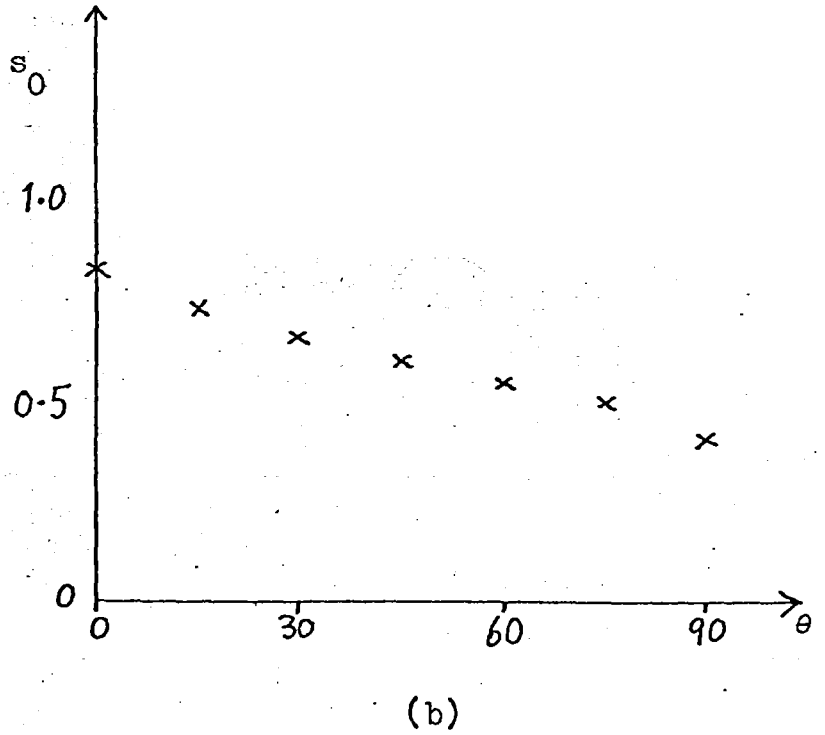


Fig.6.6. Plot of C_n against N for the three methods as in Fig.6.3.

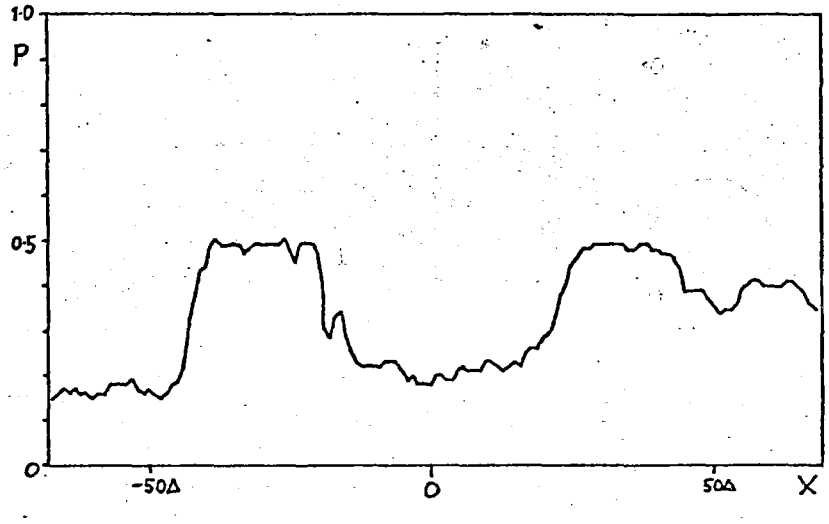


(a)

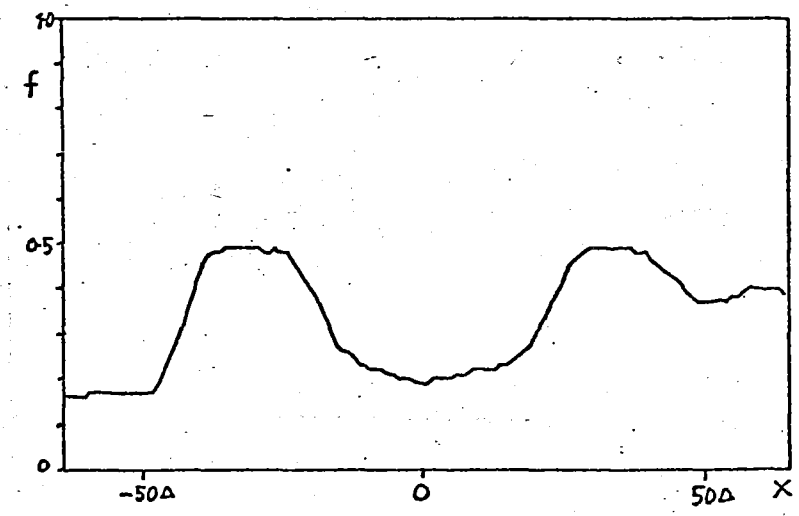


(b)

Fig.6.7. (a) C_n and (b) s_0 against θ for the psf in Fig.6.2(a) and for N = 10.

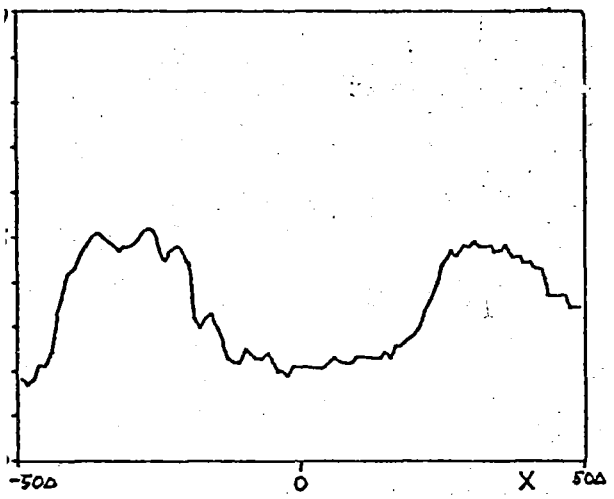


(a)

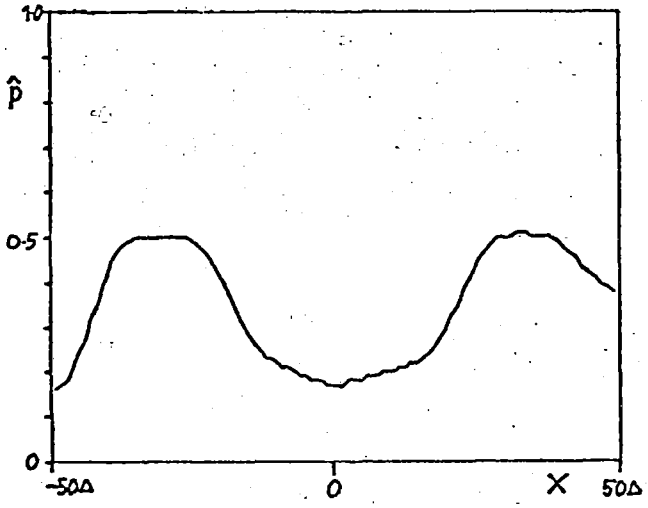


(b)

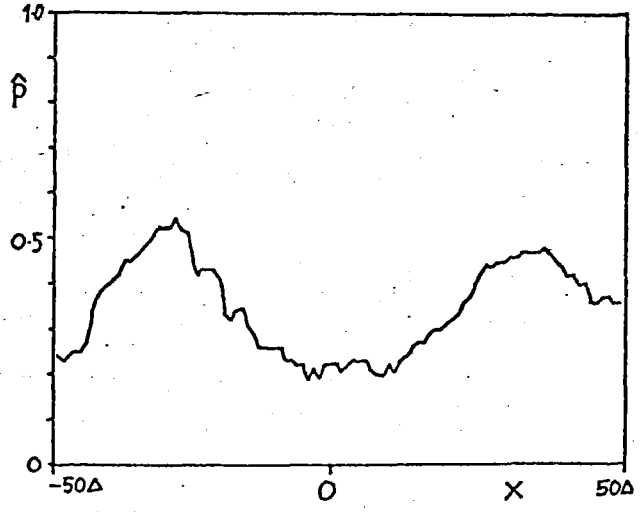
Fig.6.8 continued on next page.



(c)

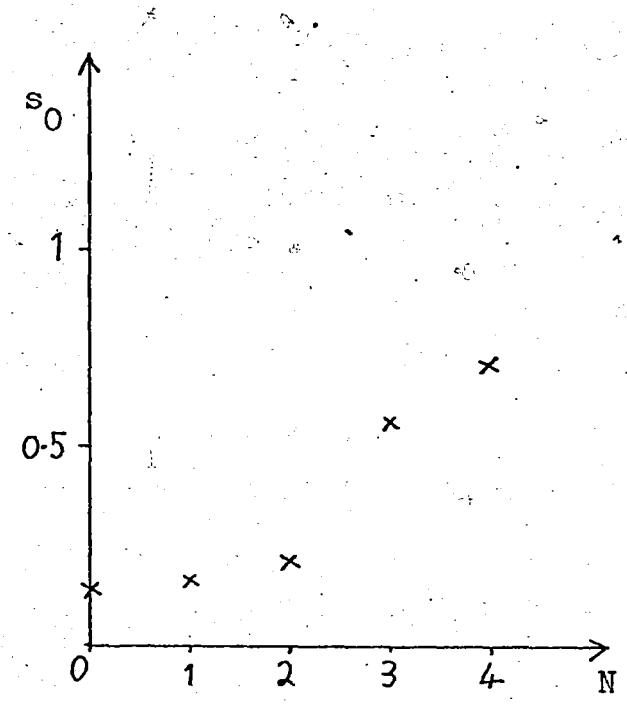


(d)

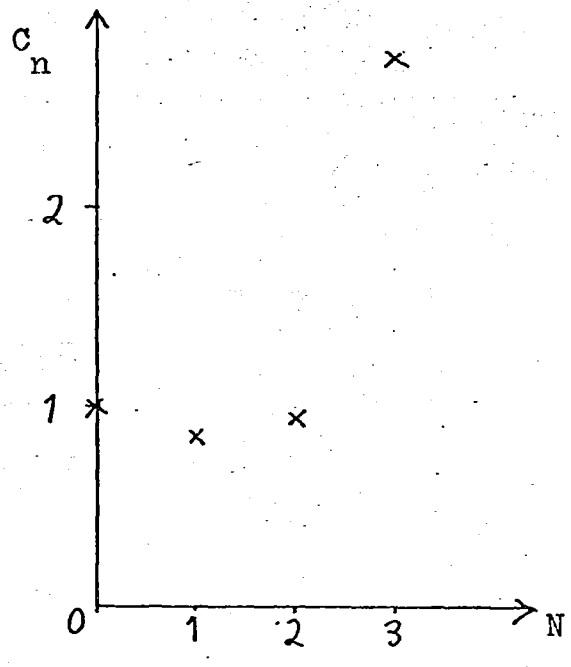


(e)

Fig. 6.8. Application to a line of an image. (a) p, (b) f; $\hat{p}$ using (c) our method, (d) method of Saleh, (e) method of Arguello et al.



(a)



(b)

Fig.6.9. (a) s_0 and (b) C_n against N for a 2-D psf which is a disc of radius 4 , and for $\theta = 0$.

Table 6.1. $\underset{\sim}{W}$, for N = 1,10, corresponding to the psf in Fig.6.2(a). As $\underset{\sim}{W}$ is symmetrical, the w_j are only given for j = 0,..,N.

N	w_j, j = 0,..,N
1	0.0000 0.5303
2	0.0000 0.0000 0.5669
3	0.0000 0.0000 0.0000 0.6124
4	0.0000 0.0000 0.0000 0.0000 0.6708
5	0.0000 0.0000 0.0000 0.0000 1.2481 -1.1094
6	0.0000 0.0000 0.0000 0.0500 1.2000 -1.2000 0.1000
7	0.0000 0.0000 0.0544 0.0000 1.1978 -1.1978 0.0000 0.1089
8	0.0000 0.0598 0.0000 0.0000 1.1952 -1.1952 0.0000 0.0000 0.1195
9	0.1324 0.0000 0.0000 0.0000 1.1921 -1.1921 0.0000 0.0000 0.0000 0.1324
10	0.0000 0.0700 0.0000 0.0000 1.1902 -1.1902 0.0000 0.0000 0.0350 0.0000 0.1050

Table 6.2. s, for N = 1,10, corresponding to the psf in Fig.6.2(a) and the $\underset{\sim}{W}$ in Table 6.1. As s is symmetrical the s_j are only given for j = 0,..,N+4.

N	s_j, j = 0,..,N+4						
1	0.3535	0.3535	0.3535	0.3535	0.1768	0.1768	
2	0.3780	0.3780	0.3780	0.1890	0.1890	0.1890	0.1890
3	0.4082	0.4082	0.2041	0.2041	0.2041	0.2041	0.2041
	0.2041						
4	0.4472	0.2236	0.2236	0.2236	0.2236	0.2236	0.2236
	0.2236	0.2236					
5	0.8320	0.0462	0.0462	0.0462	0.0462	0.0462	0.0462
	0.0462	0.0462	-0.3698				
6	0.8333	0.0333	0.0500	0.0500	0.0500	0.0500	0.0500
	0.0500	0.0333	-0.3667				
7	0.8349	0.0363	0.0363	0.0544	0.0544	0.0544	0.0544
	0.0363	0.0363	-0.3630	0.0363	0.0363		
8	0.8367	0.0398	0.0398	0.0398	0.0598	0.0598	0.0398
	0.0398	0.0398	-0.3586	0.0398	0.0398	0.0398	
9	0.8389	0.0441	0.0441	0.0441	0.0441	0.0441	0.0441
	0.0441	0.0441	-0.3532	0.0441	0.0441	0.0441	0.0441
10	0.8402	0.0467	0.0467	0.0467	0.0350	0.0350	0.0467
	0.0467	0.0467	-0.3500	0.0467	0.0467	0.0467	0.0350
	0.0350						

CHAPTER SEVEN

A SAMPLING FUNCTION APPROPRIATE FOR DECONVOLUTION

7.1 INTRODUCTION

The contents of this chapter have been submitted as a paper (McDonnell 1975b) to IEEE Transactions on Information Theory. This paper is repeated here in essentially unaltered form so that this chapter may be complete in itself.

We present a sampling theory in one and two dimensions which is appropriate for deconvolution problems. Our concern is with the possibility of sampling at a rate lower than the Nyquist rate. We introduce the concept of a "line-segment-limited" function and show that it is possible to define one- and two-dimensional sampling functions corresponding to this concept. Our viewpoint is that of image restoration. We consider the problem of obtaining, from samples of a continuous degraded image, an acceptable sampled restored image, on the understanding that the point spread function (psf) characterising the degradation is known. We show, firstly, that these new sampling functions overcome the "picket fence" effect associated with the sinc sampling function. Secondly, we define a psuedo noise level which is introduced by the approximation that the psf is "line-segment-limited". It is this pseudo noise level, and not aliasing errors, that restricts the choice of a sampling

rate less than the Nyquist rate. Thirdly, we present new criteria which allow us to choose sampling rates considerably less than the Nyquist rate.

7.2 PRELIMINARY DISCUSSION

Conventional sampling theory has been concerned with deriving conditions under which a continuous function may be reconstructed exactly from samples of the function. For example, it is well known (Goodman 1968) that a bandlimited function may be reconstructed from samples of the function taken at the Nyquist rate. In line with this result, the general approach to the problem of image restoration (Frieden 1974) has been to sample the degraded image at the Nyquist rate, and from these samples to find the best possible continuous restored image. However, there exist applications (e.g. image display using a television system) where the final interest lies in the samples themselves and there is no need to attempt to reconstruct the continuous restored image by convolution with the appropriate sampling function. Also, it is often desirable, and for practical reasons it can be necessary, to sample at a rate considerably less than the Nyquist rate. In these situations it is possible to improve upon the results of conventional sampling theory. We show how to do this in the context of image restoration.

In 7.3 we consider the one-dimensional problem of obtaining from samples of a continuous degraded image, an acceptable sampled restored image. We define what we mean by a "line-segment-limited" function and show that we may

define a corresponding sampling function. We then examine the advantages and disadvantages of using this sampling function in image restoration. We have already briefly discussed the one-dimensional problem in 6.5. In 7.3 we present a fuller discussion before showing in 7.4 that the theory can be extended to the two-dimensional case. After introducing an appropriate definition of a two-dimensional line-segment-limited function, we show that there exists a corresponding sampling function. We give our conclusions in 7.5 as well as discussing possible applications of our results.

7.3 ONE-DIMENSIONAL THEORY

We represent images by real functions of Cartesian co-ordinates x and y. The corresponding co-ordinates in the Fourier transform (FT) domain are u and v. In one dimension we denote by $f = f(x)$ the degraded image formed by the convolution of the original image $p = p(x)$ with the point spread function (psf) $h = h(x)$. Thus

$$f = f(x) = \int_{-\infty}^{\infty} p(\alpha)h(x-\alpha)d\alpha \qquad (7.1)$$

$$= p \otimes h, \qquad (7.2)$$

where $\otimes$ denotes convolution. The FT of an image is denoted by the upper case form of the lower case English letter which denotes the image, e.g.

$$F = F(u) = \int_{-\infty}^{\infty} f(x)\exp(i2\pi ux)dx. \qquad (7.3)$$

We consider f to be sampled at a rate $1/\Delta$ to give 2M+1 samples f_j, where

$$f_j = f(j\Delta), \qquad |j| \leq M. \tag{7.4}$$

The number of samples is chosen to be odd for convenience.

We assume that we are given h and the f_j. We let N_f by the Nyquist sampling rate equal to the reciprocal of the actual (or effective) extent of F. Although we make no initial assumptions about the relative sizes of N_f and $1/\Delta$, we are interested in the possibility of having

$$1/\Delta_a < 1/\Delta < N_f, \tag{7.5}$$

where $1/\Delta_a$ is the sampling rate required to display adequately the detail of interest in p. We let $\hat{p}$ denote an estimate of p. Our aim is to find acceptable values for the 2N+1 samples $\hat{p}_k$ of $\hat{p}$, where

$$\hat{p}_k = \hat{p}(k\Delta), \qquad |k| \leq N. \tag{7.6}$$

Conventionally, a continuous approach is taken to this problem. Initially f is approximated by a function $f_B = f_B(x)$, which is bandlimited to $1/2\Delta$, i.e.

$$F_B(u) = 0, \qquad |u| > 1/2\Delta, \tag{7.7}$$

where

$$f_B(x) = \sum_{j=-M}^{M} f_j \, \mathrm{sinc}(j - x/\Delta). \tag{7.8}$$

Here $\mathrm{sinc}(x) = \sin(\pi x)/\pi x$. Similarly, h is approximated by

$h_B = h_B(x)$ where

$$\left.\begin{aligned} H_B(u) &= 0, \qquad |u| > 1/2\Delta, \\ &= H(u), \quad |u| \leqslant 1/2\Delta. \end{aligned}\right\} \tag{7.9}$$

From (7.2) we have

$$P = F/H, \tag{7.10}$$

so that we may use

$$\hat{P} = F_B/H_B, \qquad |u| \leqslant 1/2\Delta \tag{7.11}$$

to obtain $\hat{P}$ and hence $\hat{p}$. Ideally

$$F_B = F, \qquad |u| \leqslant 1/2\Delta. \tag{7.12}$$

However, this is only true if $N_f \leqslant 1/\Delta$. If it is true, (7.11) yields

$$\hat{p}(x) = p(x) \otimes \mathrm{sinc}(x/\Delta). \tag{7.13}$$

The $\hat{p}_k$ are then found by sampling $\hat{p}$. If

$$p(x) = \delta(x-j\Delta-\alpha\Delta), \quad 0<\alpha<1, \quad -N<j<N-1, \tag{7.14}$$

we obtain

$$\hat{p}_k = \mathrm{sinc}(k-j-\alpha), \qquad |k| \leqslant N. \tag{7.15}$$

Now, (7.15) exhibits the "picket fence" effect (Bergland 1969), i.e. a delta function off a sampling point in the original image has an effect at every sampling point in the restored image. This is of no consequence if p is band-

limited to $1/2\Delta$. However, if p is not bandlimited to $1/2\Delta$, the $\hat{p}_k$ given by sampling (7.13) may be unacceptably different, in the mean square error sense, from the result of directly sampling p. If (7.5) holds, (7.12) may be considerably in error because of aliasing. This will cause errors in $\hat{P}$ in (7.11), and hence in the $\hat{p}_k$ given by sampling (7.13).

We now consider our alternative approach. First we define a one-dimensional sampling function $s_1 = s_1(x)$ by

$$\left.\begin{aligned} s_1(x) &= 0, && |x| > \Delta, \\ &= 1 - |x|/\Delta, && |x| \leqslant \Delta. \end{aligned}\right\} \tag{7.16}$$

We consider h to have a finite number 2J+1 of samples h_j, where M = J+N and

$$h_j = h(j\Delta), \qquad |j| \leqslant J. \tag{7.17}$$

We define the sampled psf $h_s = h_s(x)$ by

$$h_s(x) = \sum_{j=-J}^{J} h_j \, \delta(x - j\Delta). \tag{7.18}$$

We now define $h_c = h_c(x)$ by

$$h_c = h_s \otimes s_1. \tag{7.19}$$

Here h_c is the continuous function obtained by linearly interpolating between adjacent samples of h. We say that h is line-segment-limited if

$$h = h_c. \tag{7.20}$$

From (7.19) we see that s_1 is the sampling function corresponding to the concept of a line-segment-limited function. We now approximate h by h_c.

We have shown in 6.5 that $h_s(x-j\Delta)$, where $|j| \leqslant N$, may be ideally restored to a delta function by convolution with an ideal discrete filter array $\underset{\sim}{W}_I$, where

$$\delta(x - j\Delta) = \underset{\sim}{W}_I \otimes h_s(x - j\Delta). \qquad (7.21)$$

If the discrete convolution in (7.21) is to be carried out using the fast Fourier transform (FFT) algorithm (Bergland 1969), it is easy to find the $\underset{\sim}{W}_I$ such that (7.21) holds. This corresponds to direct inverse filtering using the FFT. However, in this case, convolution in (7.21) must be replaced by circular convolution. This is equivalent to assuming that $\underset{\sim}{W}_I$ is periodic and that the convolution is to be evaluated over one period of $\underset{\sim}{W}_I$. We formulate our theory for the case of convolution with a discrete filter array as in (7.21). However, our results are also applicable to the case of circular convolution using the FFT.

From (7.21) we see that if $h_c(x-j\Delta)$ is sampled to give $h_s(x-j\Delta)$ and convolved with $\underset{\sim}{W}_I$ the result is $\delta(x-j\Delta)$. However, in general, h_c is not centred on a sampling point. If p is given by (7.14), then

$$f(x) = h_c(x-j\Delta-\alpha\Delta). \qquad (7.22)$$

The special property of h_c noted in 6.5 is that $h_c(x-j\Delta-\alpha\Delta)$ has the same sampled values as

$$(1 - \alpha)h_c(x - j\Delta) + \alpha h_c(x - (j+1)\Delta). \qquad (7.23)$$

Consequently, when $h_c(x-j\Delta-\alpha\Delta)$ is sampled and convolved with $\underset{\sim}{W}_I$ we obtain

$$\left.\begin{aligned} \hat{p}_j &= 1-\alpha, \\ \hat{p}_{j+1} &= \alpha, \\ \text{and} \qquad \hat{p}_k &= 0, \quad k \neq j \text{ or } j+1, \quad |k| \leq N. \end{aligned}\right\} \tag{7.24}$$

On comparing (7.24) with (7.15) we see that we have overcome the picket fence effect. From (7.16) and Fig. 7.1 we see that the samples $\hat{p}_j$ and $\hat{p}_{j+1}$ in (7.24) are equal to the non-zero samples of $s_1(x-j\Delta-\alpha\Delta)$. Thus, when p is any delta function given by (7.14), the $\hat{p}_k, |k| \leq N$, are samples of

$$\hat{p} = p \otimes s_1 . \tag{7.25}$$

Our theory is linear so that for a general p, the $\hat{p}_k$ are samples of (7.25). The samples $\hat{p}_k$ are obtained by convolving the f_j with ideal filter array $\underset{\sim}{W}_I$. We note that (7.25) corresponds to (7.13) in the previous approach, and that the sum of the $\hat{p}_k$ in 7.24 is one as required for normalisation.

Because the width of s_1 is only 2Δ, the ideal sampled restored image obtained from (7.25) is similar to the result of directly sampling p. Indeed, in our opinion it is even better than the result of directly sampling p, since information at any point off a sampling point is retained at the adjacent sampling points in inverse proportion to its distance from them.

The most important aspect of (7.25) is that if h is

line-segment-limited, then (7.25) holds no matter how large N_p is, or how inaccurate (7.12) is. In other words, at any given sampling rate $1/\Delta$, our sampled restored image is unaffected by aliasing if h is line-segment-limited. This indicates that aliasing errors do not restrict the choice of a sampling rate less than the Nyquist rate.

In practice h is not line-segment-limited. However, most practical one-dimensional psf's may be approximated by a line-segment-limited psf at a sampling rate satisfying (7.5). This approximation, $h = h_c$, may be considered to introduce a psuedo noise level

$$\sigma_h = \frac{1}{\sqrt{2J+1}} \int_{-\infty}^{\infty} |h(x) - h_c(x)|\,dx \tag{7.26}$$

on f.

In practice, a filter array $\underset{\sim}{W} = \{w_k; |k| \leq K\}$ is used rather than the ideal filter array $\underset{\sim}{W}_I$. The samples of the restored image are thus

$$\hat{p}_k = \sum_{n=-K}^{K} w_n f_{k-n}, \quad |k| \leq N, \quad |k-n| \leq M. \tag{7.27}$$

Use of the filter array $\underset{\sim}{W}$ gives rise to a noise amplification which we denote by C_n.

Our experience is that a certain amount of a priori information is almost always available for use in the restoration procedure. This information includes an estimate not only of the psf, but also of $1/\Delta_a$, of the maximum noise amplification C_n^{max} which $\underset{\sim}{W}$ can be allowed to generate, and

of the maximum acceptable noise level σ_n^{max} on the $\hat{p}_k$. We feel that, in the context of image restoration by discrete convolution with a finite filter array, or by discrete circular convolution using the FFT, the following criteria lead to a sampling rate $1/\Delta$ more suitable than the Nyquist rate. Firstly, (7.5) should be satisfied. Secondly, the noise level introduced by the approximation (7.20) should not, when convolved with $\underset{\sim}{W}$, give rise to a noise level on the $\hat{p}_k$ which is greater than σ_n^{max}. In other words, the sampling rate $1/\Delta$ should be the slowest rate satisfying

$$1/\Delta_a < 1/\Delta < N_f \tag{7.5}$$

and

$$\sigma_h \, c_n^{max} < \sigma_n^{max} . \tag{7.28}$$

7.4 TWO-DIMENSIONAL THEORY

We now show how the theory we developed in the previous section may be extended to cover the case when the psf is two-dimensional. We let $f = f(x,y)$, $h = h(x,y)$, $p = p(x,y)$ and $\hat{p} = \hat{p}(x,y)$. We assume that sampling is carried out on a rectangular grid and that, for simplicity, the sampling rate is $1/\Delta$ in both the x and y directions. We also assume that the number of samples of any function is the same in both the x and y directions.

Corresponding to (7.4) and (7.6), we define samples $f_{j,k}$ of f and $\hat{p}_{j,k}$ of $\hat{p}$ by:

$$f_{j,k} = f(j\Delta,k\Delta); \quad |j|, |k| \leq M, \tag{7.29}$$

$$\hat{p}_{j,k} = p(j\Delta,k\Delta); \quad |j|, |k| \leq N. \tag{7.30}$$

We assume that the Nyquist sampling rate corresponding to f in each of the x and y directions is N_f.

We replace s_1 in (7.16) by a two-dimensional sampling function $s_2 = s_2(x,y)$, illustrated in Fig. 7.2, and defined by

$$s_2(x,y) = s_1(x)\, s_1(y). \tag{7.31}$$

Corresponding to (7.17) we define samples $h_{j,k}$ of h by

$$h_{j,k} = h(j\Delta,k\Delta); \quad |j|, |k| \leq J. \tag{7.32}$$

We define the sampled psf $h_s = h_s(x,y)$ by

$$h_s(x,y) = \sum_{j=-J}^{J} \sum_{k=-J}^{J} h_{j,k}\delta(x-j\Delta)\delta(y-k\Delta), \tag{7.33}$$

and $h_c = h_c(x,y)$ by

$$h_c = h_s \otimes s_2. \tag{7.34}$$

Here h_c is the continuous function obtained by first linearly interpolating between adjacent samples of h to give h_c along the lines

$$\left.\begin{aligned} x &= j\Delta, \quad |j| \leq J, \\ y &= k\Delta, \quad |k| \leq J. \end{aligned}\right\} \tag{7.35}$$

Then h_c is found at points off these lines by linearly interpolating in a direction parallel to the x or y axes. We say that h is line-segment-limited in two dimensions if

$$h = h_c. \tag{7.36}$$

From (7.34) we see that s_2 is the sampling function corresponding to the concept of a line-segment-limited function in two dimensions. We now approximate h by h_c.

Corresponding to (7.21), we define $\underset{\sim}{W}_I$ to be an ideal two-dimensional discrete filter array such that

$$\delta(x-j\Delta)\ \delta(y-k\Delta) = \underset{\sim}{W}_I \otimes h_s(x-j\Delta,y-k\Delta). \tag{7.37}$$

We now consider the case when

$$p(x,y) = \delta(x-j\Delta-\alpha\Delta)\ \delta(y-k\Delta-\beta\Delta);$$

$$0 < \alpha < 1,\ 0 < \beta < 1,\ -N < j < N-1,\ -N < k < N-1. \tag{7.38}$$

This gives

$$f(x,y) = h_c(x-j\Delta-\alpha\Delta,y-k\Delta-\beta\Delta). \tag{7.39}$$

Corresponding to (7.23), $h_c(x-j\Delta-\alpha\Delta,y-k\Delta-\beta\Delta)$ has the same sampled values as

$$(1-\alpha)(1-\beta)h_c(x-j\Delta,y-k\Delta) + \alpha(1-\beta)h_c(x-(j+1)\Delta,y-k\Delta) +$$

$$(1-\alpha)\beta h_c(x-j\Delta,y-(k+1)\Delta) + \alpha\beta h_c(x-(j+1)\Delta,y-(k+1)\Delta). \tag{7.40}$$

Consequently, when f is given by (7.39) is sampled and convolved with $\underset{\sim}{W}_I$, we obtain

$$\left.\begin{aligned} \hat{p}_{j,k} &= (1-\alpha)\ (1-\beta), \\ \hat{p}_{j+1,k} &= \alpha(1-\beta), \\ \hat{p}_{j,k+1} &= (1-\alpha)\beta, \\ \hat{p}_{j+1,k+1} &= \alpha\beta, \end{aligned}\right\} \qquad (7.41)$$

and $\hat{p}_{l,m} = 0;\ l \neq j \text{ or } j+1;\ m \neq k \text{ or } k+1;\ |l|, |m| \leq N.$

We now note that in (7.41) the samples $\hat{p}_{l,m}$ are samples of $s_2(x-j\Delta-\alpha\Delta, y-k\Delta-\beta\Delta)$, and that the sum of the samples is one as required for normalisation. We thus obtain the result corresponding to (7.25), that for a general p the $\hat{p}_{l,m}$ are samples of

$$\hat{p} = p \otimes s_2. \qquad (7.42)$$

These samples $\hat{p}_{l,m}$ are obtained by convolving the $f_{j,k}$ with the ideal filter array $\underset{\sim}{W}_I$.

Our discussion of the one-dimensional case applies also in two dimensions, except that in place of (7.26) we have

$$\sigma_h = \frac{1}{2J+1} \iint_{-\infty}^{\infty} |h(x,y) - h_c(x,y)|\ dxdy. \qquad (7.43)$$

Also the practical filter array $\underset{\sim}{W}$ becomes $\underset{\sim}{W} = \{w_{j,k}; |j|, |k| \leq K\}$ so that the samples of the restored image are

$$\hat{p}_{j,k} = \sum_{l=-K}^{K} \sum_{m=-K}^{K} w_{l,m}\ f_{j-l,k-m}; |j|, |k| \leq N; |j-l|, |k-m| \leq M. \qquad (7.44)$$

7.5 CONCLUSIONS

We have presented a sampling theory in one and two dimensions which is appropriate for deconvolution problems. This theory is based on the properties of line-segment-limited functions and their corresponding sampling functions. We have defined a psuedo noise level which is introduced by the approximation that the psf is line-segment-limited. It is this psuedo noise level, and not aliasing errors, that restricts the choice of a sampling rate less than the Nyquist rate. We have introduced new sampling criteria which can lead to the choice of a sampling rate considerably less than the Nyquist rate. We have also shown that our new sampling functions overcome the "picket fence" effect associated with the sinc sampling function. Our ideas can be applied directly to deconvolution problems.

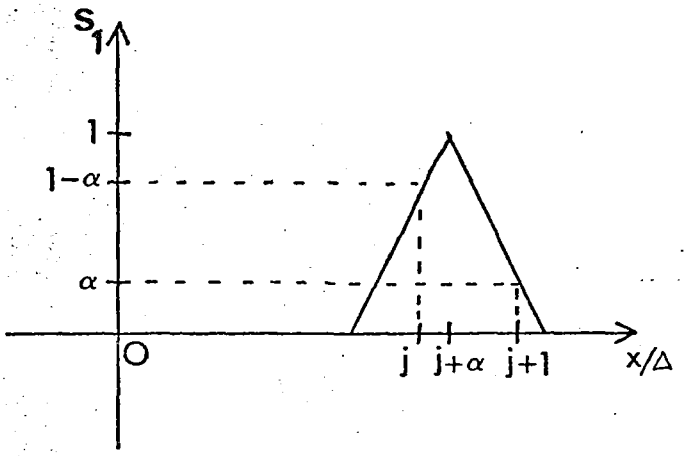


Fig. 7.1 $s_1(x-j\Delta-\alpha\Delta)$.

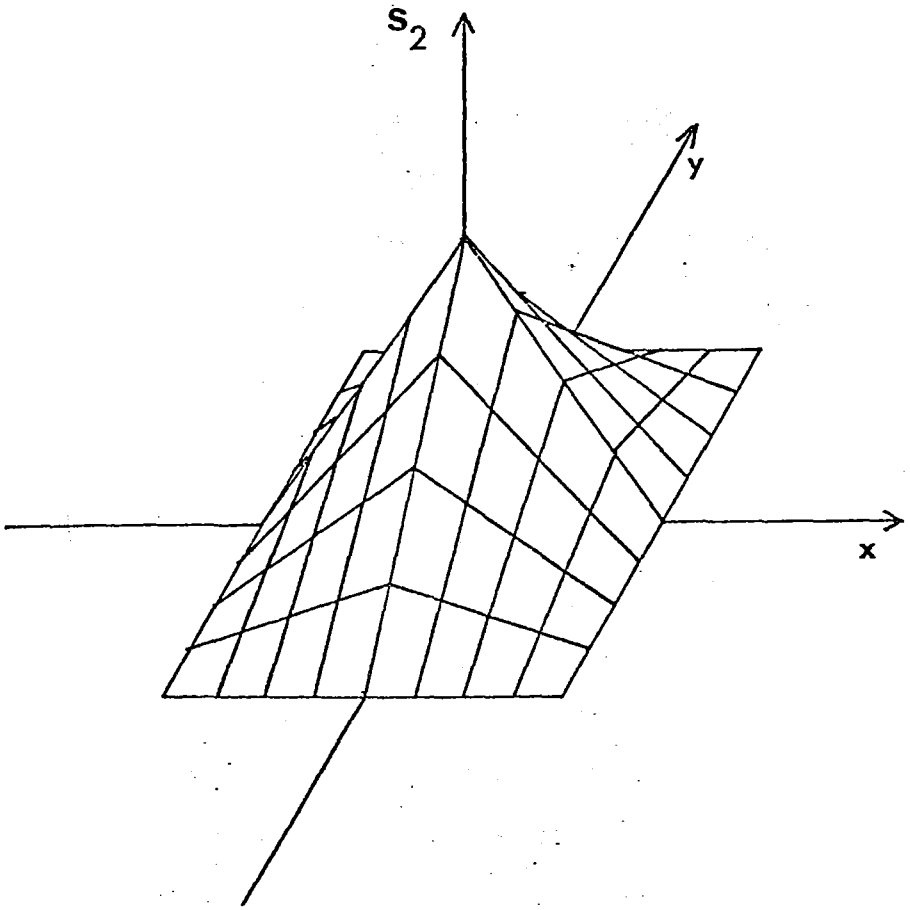


Fig. 7.2 $s_2(x,y)$.

CHAPTER EIGHT

CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

8.1 CONCLUDING REMARKS

In this thesis a number of soundly based techniques are presented for the digital restoration of optically degraded images. Degraded images are divided into the classes $\mathcal{G}$ and $\mathcal{S}$ - those that are and are not respectively truncated by their recording frames. It is shown that conventional restoration techniques work well for images of class $\mathcal{S}$, but are unsoundly based for images of class $\mathcal{G}$. The new techniques presented here are shown to be practical and efficient through the use of computer simulations and actual examples. The most useful of these techniques are the edge extension method introduced in chapter 4 and the improved nonrecursive digital filtering method in chapter 6. The edge extension method is best suited to the a posteriori restoration of an image which is stored in a computer in sampled form. The nonrecursive method is best suited to real-time digital implementation where its use may lead to a significant improvement in the resolution of electron microscope images as is discussed in the next section. The other restoration techniques presented in chapters 3, 4 and 5, although practical, are restricted in their application compared to the above two methods. They are, nevertheless, technically interesting. It is hoped that

the ideas on sampling theory developed in chapter 7 will prove useful for choosing sampling rates in many areas. Although significant progress in image restoration is reported here, this thesis has raised more questions than it has answered. Promising areas for further research are discussed in the next section.

8.2 SUGGESTIONS FOR FURTHER RESEARCH

Little attempt has been made to restore images of the earth and other planets obtained from satellites, although considerable enhancement of such images has been carried out. The reason for this is that at the sampling rates used to transmit data from e.g. the ERTS satellites, the quality of the data is such that restoration is unnecessary. It would not be difficult to increase the rate at which the data is sampled. If this was done, it would be feasible to restore the resulting class $\mathcal{G}$ images by the methods presented here. This could be done in real-time using a 2-D nonrecursive digital filter (discussed in more detail later in the context of electron microscopy). Alternatively, selected regions of interest of the recorded image could be restored using the edge extension method. These approaches would lead to improved resolution in satellite images.

Perhaps the most promising applications of image restoration lie in the area of high resolution electron-microscopy. It is well known that amplitude contrast and phase contrast electron microscope images may be modelled as the convolution of an original image with a psf which is

approximately SI over any small region of interest. The psf may be deduced from the instrument settings and the calibration constants of the microscope. Therefore, the class $\mathcal{G}$ degraded images may be restored a posteriori using the edge extension method. However, this involves considerable computation for each restored image. A more attractive scheme would be to carry out real-time 2-D nonrecursive digital filtering. This could be accomplished as follows. The electron microscope image would be displayed as a television picture. Scanning electron microscope images are already displayed in this form, and it is technically feasible to display transmission electron microscope images in this way also. When restoration of a high resolution image is required, it would be done in real-time using a 2-D nonrecursive digital filter incorporated into the television scanning system. Time delays equivalent to several lines of the image would be required. The filter components would be calculated and supplied by a small on-line computer. This computer would, in turn, be supplied interactively with data from the operator by means of the control switches on the electron microscope. The computer would only be used to update the filter elements when required.

In this thesis, the deconvolution problem has been dealt with mainly in the context of image restoration. The image restoration techniques derived here are also applicable to 1-D signal processing problems and may prove useful in this area as well.

In 1.3.3 an image segmentation technique due to Stockham et al (1975) is discussed. This technique could be improved as follows by the use of the edge extension method. The method of Stockham et al (1975) would be used to obtain an initial estimate of $|H|$ from (1.18). If h is of a simple form (cf. Table 1.1), H and hence h could be deduced from $|H|$. It would then be possible to find the T corresponding to h. The edge extension method could then be used to find an overlapped extended image $\hat{e}^j$ corresponding to f^j for each image segment. An improved estimate of $|H|$, and hence h, could then be obtained using (1.18) with F^j replaced by $\hat{E}^j$.

The effect of noise on a degraded image is to modify the amplitude and phase of the frequency components of the image. When the degraded image is inverse filtered, the amplitude of the noise frequency components tends to be increased. Filters such as the Wiener filter are designed to minimise this amplification. However, no attempt is made to correct the phase errors caused by the noise. These phase errors can be large at high frequencies. In my experience the phases of the Fourier coefficients of an image are highly correlated. It seems likely that knowledge of the general correlation between the phases of frequency components of an image could be used to reduce phase errors.

In chapter 6 the optimum nonrecursive filter is derived for any given number of filter elements, assuming that the filter elements are uniformly spaced. An interesting problem yet to be solved is as follows: for a given

number of filter elements what are the best spacings and values for the filter components?

A considerable amount of work needs to be done on the sampling ideas presented in chapter 7, before they can become widely useful. In particular, the noise correlation needs to be included in the theory in chapter 7, and the noise level introduced by aliasing errors needs to be more clearly understood.

Finally, the quality of the restored images presented in this thesis is limited more by by the display devices used, than by the restoration techniques themselves. The images produced by the electrostatic lineprinter and the teletype lineprinter are of low quality. A cheap, efficient method is required for displaying computer generated images of good tonal quality.

REFERENCES

Adler, M., Andrews, H.C. 1974 "Space variant point spread function pseudoinversion," in "Image processing research," USCIPI Report 530 (Pratt, W.K. Ed.).

Agarwal, R.C., Burrus, C.S. 1974 "Fast convolution using Fermat number transforms with applications to digital filtering," IEEE Trans. Acoustics, Speech and Signal Processing ASSP-22, 87.

Anderson, G.B., Huang, T.S. 1971. "Frequency-domain image errors," Pattern recognition 3, 185.

Andrews, H.C. 1970. "Computer techniques for image processing," New York: Academic Press.

Andrews, H.C. 1972 "N Topics in Search of an Editorial: Heuristics, Superresolution and Bibliography," Proc. IEEE 7, 891.

Andrews, H.C. 1974 "Digital image restoration - a survey," Computer 7(5), 36.

Andrews, H.C., Tescher, A.G., Kruger, R.P. 1972 "Image processing by digital computer," IEEE Spectrum 9, 20.

Arquello, R.J., Sellner, H.R., Stuller, J.A. 1972 "Transfer function compensation of sampled imagery," IEEE Trans. Comput. C-21, 812.

Backus, G., Gilbert, G. 1970. "Uniqueness in the inversion of inaccurate grass earth data," Phil. Trans. Roy. Soc. London A266, 123.

Barnes, C.W. 1966 "Object restoration in a diffraction-limited imaging system," J. Opt. Soc. Am. 56, 575.

Bates, R.H.T. 1969 "Contributions to the theory of intensity interferometry," Mon. Not. R. astr. Soc. 142, 413.

Bates, R.H.T., Gough, P.T. 1975 "New outlook on processing radiation from objects viewed through randomly fluctuating media," IEEE Trans. Comput. C-24, 449.

Bates, R.H.T., Kennedy, W.K., McDonnell, M.J. 1974 "Efficient digital restoration of images blurred by linear motion," Letters in Applied and Engineering Sciences 2, 133.

Bates, R.H.T., Napier, P.J. 1972 "Identification and removal of phase errors in interferometry," Mon. Not. R. astr. Soc. 158, 405.

Bates, R.H.T., Napier, P.J., McKinnon, A.E., McDonnell, M.J. 1975 "Self-consistent deconvolution: I-theory," to be published in Optik.

Bergland, G.D. 1969 "A guided tour of the fast Fourier transform," IEEE Spectrum 6, 41.

Billingsley, F.C. 1973 "Digital image processing for information extraction," Int. J. Man-machine studies 5, 203.

Biraud, Y. 1969 "A new approach for increasing the resolving power by data processing," Astron. and Astrophys. 1, 124.

Bryngdahl, D., Lohmann, A.W. 1968 "Holographic penetration of turbulence," in Proc. NASA/ERC seminar (M. Nagel Ed.) Cambridge, Mass.

Campbell, K., Wecksung, G.W., Mansfield, C.R. 1974 "Spatial filtering by digital holography," Optical Eng. 13, 175.

Cole, E.R. 1973 "The removal of unknown image blurs by homomorphic filtering," Ph.D. Dissertation, Dept. of Electrical Engineering, Univ. of Utah, Salt Lake City.

Cooley, J.W., Tukey, J.W. 1965 "An algorithm for the machine calculation of complex Fourier series," Math. Comput. 19, 297.

De Santis, P., Gori, F. 1975 "On an iterative method for superresolution," Opt. Acta 22, 691.

Ekstrom, M.P. 1973a "A numerical algorithm for identifying spread functions of shift invariant imaging systems," IEEE Trans. Comput. C-22, 322.

Ekstrom, M.P. 1973b "A spectral characterisation of the ill-conditioning in numerical deconvolution," IEEE Trans. Audio Electroacoust. AU-21, 345.

Ekstrom, M.P. 1973c "On the numerical feasibility of digital image restoration," Proc. IEEE 61, 1155.

Erickson, H.P. 1973 "The Fourier transform of an electron micrograph - first order and second order of image formation," Advances in Optical and Electron Microscopy 5, 163.

Erickson, H.P., Klug, A. 1971 "Measurement and compensation of defocussing and aberrations by Fourier processing of electron micrographs," Phil. Trans. Roy. Soc. London B261, 105.

Frieden, B.R. 1967 "Band-unlimited reconstruction of optical objects and spectra," J. Opt. Soc. Am. 57, 1013.

Frieden, B.R. 1968 "Optimum nonlinear processing of noisy images," J. Opt. Soc. Am. 58, 1272.

Frieden, B.R. 1972 "Restoring with maximum likelihood and maximum entropy," J. Opt. Soc. Am. 62, 511.

Frieden, B.R. 1973 "Restoration of pictures by Monte Carlo allocation of "grains"," in Proc. Aug. Meeting Washington, D.C.: Opt. Soc. Amer.

Frieden, B.R. 1974 "Image restoration by discrete convolution of minimal length," J. Opt. Soc. Am. 64, 682.

Frieden, B.R., Burke, J.J. 1972 "Restoring with maximum entropy II: superresolution of photographs of diffraction-blurred images," J. Opt. Soc. Am. 62, 1207.

Gennery, D.B. 1973 "Determination of optical transfer function by inspection of frequency domain plot," J. Opt. Soc. Am. 63, 1571.

Gerchberg, R.W. 1974 "Superresolution through error energy reduction," Opt. Acta 21, 709.

Goodman, J.W. 1968 "Introduction to Fourier optics," New York: McGraw-Hill.

Gough, P.T., Bates, R.H.T. 1972 "Computer generated holograms for processing radiographic data," Computers and Biomed. Res. 5, 700.

Granger, E.M. 1968 "Restoration of images degraded by spatially varying smear," in Proc. NASA/ERC Seminar (M. Nagel Ed.) Cambridge, Mass.

Hall, E.L. 1972 "A comparison of computations for spatial frequency filtering," Proc. IEEE 59, 887.

Hall, E.L., Kruger, R.P., Dwyer, S.J., Hall, D.L., McLaren, R.W., Lodwick, G.S. 1971 "A survey of preprocessing and feature extraction techniques for radiographic images," IEEE Trans. Comput. C-20, 1032.

Harris, J.L. 1964 "Diffraction and resolving power", J. Opt. Soc. Am. 54, 931.

Harris, J.L. Sr. 1966 "Image evaluation and restoration," J. Opt. Soc. Am. 56, 569.

Harris, J.L. Sr. 1968 "Potential and limitations of techniques for processing linear motion-degraded imagery," in Proc. NASA/ERC Seminar (M. Nagel Ed.), Cambridge, Mass.

Helstrom, C.W. 1967 "Image restoration by the method of least squares," J. Opt. Soc. Am. 57, 297.

Honda, T., Tsujiuchi, J. 1975 "Restoration of linear-motion blurred pictures by image scanning method (Effect of total width of the scanning aperture)," Opt. Acta 22, 537.

Honda, T., Tsujiuchi, J., Ishiquro, S. 1974 "Image processing by multiple aperture scanning," Opt. Acta 21, 653.

Hoppe, W. 1970 "Principles of electron structure research at atomic resolution using conventional electron microscopes for the measurement of amplitudes and phases," Acta Cryst. A26, 414.

Horner, J.L. 1969 "Optical spatial filtering with the least-mean-square-error filter," J. Opt. Soc. Am. 59, 297.

Huang, T.S. 1966 "Some notes on film grain noise," in NSF Summer Study Rep. (S. Morgan Ed.), Woods Hole, Mass.

Huang, T.S., Schreiber, W.F., Tretiak, O.J. 1971 "Image processing," Proc. IEEE 59, 1586.

Hunt, B.R. 1972 "Data structures and computational organisation in digital image enhancement," Proc. IEEE 60, 884.

Hunt, B.R. 1975 "Digital Image Processing," Proc. IEEE 63, 693.

Hunt, B.R., Janney, D.H. 1974 "Digital image processing at Los Alamos scientific laboratory," IEEE Computer 7(5), 57.

Inuiya, M., Ichioka, Y. 1973 "Image restoration of blurring due to uniform linear motion by TV techniques," Opt. Commun. 8, 382.

Jansson, P.A. 1970 "Method for determining the response function of a high-resolution infrared spectrometer," J. Opt. Soc. Am. 60, 184.

Jansson, P.A., Hunt, R.M., Plyler, E.K. 1970 "Resolution enhancement of spectra," J. Opt. Soc. Am. 60, 596.

Kohler, D., Mandel, L. 1973 "Source reconstruction from the modulus of the correlation function : a practical approach to the phase problem of optical coherence theory," J. Opt. Soc. Am. 63, 126.

Labeyrie, A. 1970 "Attainment of diffraction limited resolution in large telescopes by Fourier analysing speckle patterns in star images," Astron. and Astrophys. 6, 85.

Lahart, M.J. 1974 "Maximum-likelihood restoration of nonstationary imagery," J. Opt. Soc. Am. 64, 17.

Lee, W.H. 1970 "Sampled Fourier transform hologram generated by computer," Appl. Opt. 9, 639.

Lee, S.H. 1974 "Mathematical operations by optical processing," Optical Eng. 13, 198.

Lewis, B.L., Sakrison, D.J. 1975 "Computer enhancement of scanning electron micrographs," IEEE Trans. Circuits and Systems CAS-22, 267.

Lohmann, A.W., Werlich, H.W. 1967 "Holographic production of spatial filters for code translation and image restoration," Phys. Lett. 25A, 570.

Lui, B., Gallagher, N.C. 1974 "Optimum Fourier transform division filters with magnitude constraint," J. Opt. Soc. Am. 64, 1227.

MacAdam, D.P. 1970 "Digital image restoration by constrained deconvolution," J. Opt. Soc. Am. 60, 1617.

Maréchal, A., Croce, P. 1953 "Un filtre de frequencies spatiales pour l'amelioration du contraste des images optiques," C. R. Acad. Sci. 127, 607.

Mascarenhas, N.D.A., Pratt, W.K. 1975 "Digital image restoration under a regression model," IEEE Trans. Circuits and Systems CAS-22, 252.

McDonnell, M.J. 1975a "Nonrecursive image restoration using a finite filter array," Optik 43, 159.

McDonnell, M.J. 1975b "A sampling function appropriate for deconvolution," to be published in IEEE Trans. on Information Theory.

McDonnell, M.J., Bates, R.H.T. 1975a "Preprocessing of degraded images to augment existing restoration methods," Computer Graphics and Image Processing 4, 25.

McDonnell, M.J., Bates, R.H.T. 1975b "Restoring parts of scenes from blurred photographs," Optics Commun. 13, 347.

McDonnell, M.J., Kennedy, W.K., Bates, R.H.T. 1975 "Identifying and overcoming practical problems of digital image restoration," submitted to the New Zealand Journal of Science.

McGlamery, B.L. 1967 "Restoration of turbulence-degraded images," J. Opt. Soc. Am. 57, 293.

McKinnon, A.E., McDonnell, M.J., Napier, P.J., Bates, R.H.T. 1975 "Self-consistent deconvolution; II-applications," To be published in Optik.

Mersereau, R.M., Oppenheim, A.V. 1974 "Digital reconstruction of multidimensional signals from their projections," Proc. IEEE 62, 1319.

Meuller, D.F., Reynolds, G.O. 1967 "Image restoration by the removal of random media degradations," J. Opt. Soc. Am. 57, 11.

Nagel, M. (Ed.) 1968 "Evaluation of motion-degraded images," in Proc. NASA/ERC Seminar, Cambridge, Mass.

Nathan, R. 1971 "Image processing for electron microscopy, I: Enhancement procedures," Advances in Optical and Electron Microscopy 4, 85.

Oppenheim, A.V., Schafer, R.W., Stockham, T.G. 1968 "Nonlinear filtering of multiplied and convolved signals," Proc. IEEE 56, 1264.

Patterson, C.L., Buechler, G. 1974 "Digital image processing at the aerospace corporation," IEEE Computer 7(5), 5.

Peters, T.M. 1973 "Image reconstruction from projections," Ph.D. Dissertation, Dept. of Electrical Engineering, Univ. of Canterbury, Christchurch, New Zealand.

Philip, J. 1963 "Reconstruction from measurements of positive quantities by the maximum-likelihood method," J. Math. Analysis and Appl. 7, 327.

Philip, J. 1973 "Restoration of pictures by quadratic programming and by Fourier transformation in the complex domain," Proc. IEEE 61, 468.

Phillips, D.L. 1962 "A technique for the numerical solution of certain integral equations of the first kind," J. ACM 9, 84.

Rabiner, R.L. 1971 "Techniques for designing finite-duration impulse-response digital filters," IEEE Trans. Commun. Technol. COM-19, 188.

Rabiner, R.L., Rader, C.M. (Eds) 1972 "Digital Signal Processing," IEEE Press Selected Reprint Series.

Richardson, W.H. 1972 "Bayesian-based iterative method of image restoration," J. Opt. Soc. Am. 62, 55.

Riemer, T.E., McGillem, C.D. 1973 "Constrained optimisation of image restoration," Appl. Opt. 12, 2027.

Rino, C.L. 1969 "Bandlimited image restoration by linear mean-square estimation," J. Opt. Soc. Am. 59, 547.

Robbins, G.M. 1970 "Image restoration for a class of linear spatially-variant degradations," Pattern Recognition 2, 91.

Robbins, G.M., Huang, T.S. 1972 "Inverse filtering for linear shift-variant imaging systems," Proc. IEEE 60, 862.

Robinson, G.S. 1972 "Logical convolution and discrete Walsh and Fourier power spectra," IEEE Trans. Audio and Electroacoustics AU-20, 271.

Roetling, P.G., Haas, R.C., Kinzly, R.E. 1968 "Some practical aspects of measurement and restoration of motion-degraded images," in Proc. NASA/ERC Seminar (M. Nagel Ed.), Cambridge, Mass.

Rosenfeld, A. 1969 "Picture Processing by Computer," New York: Academic Press.

Rosenfeld, A. 1972 "Picture processing," Computer Graphics and Image Processing 1, 394.

Rushforth, C.K., Harris, R.W. 1968 "Restoration, resolution and noise," J. Opt. Soc. Am. 58, 539.

Saleh, B.E.A. 1974 "Trade-off between resolution and noise in restoration by superposition of images," Appl. Opt. 13, 1833.

Sawchuk, A.A. 1973 "Space-variant system analysis of image motion," J. Opt. Soc. Am. 63, 1052.

Saxton, W.D. 1974 "A new computer language for electron image processing," Computer Graphics and Image Processing 3, 266.

Shanks, J.L. 1970 "The design of stable two-dimensional recursive filters," in Proc. Kelly Communications Conf. (Univ. Missouri, Rolla).

Silverman, H.S., Pearson, A.E. 1973 "On deconvolution using the discrete Fourier transform," IEEE Trans. Audio Electroacoust. AU-21, 112.

Slepian, D. 1967a "Linear least-squares filtering of distorted images," J. Opt. Soc. Am. 57, 918.

Slepian, D. 1967b "Restoration of photographs blurred by image motion," Bell Syst. Tech. J. 46, 2353.

Slepian, D., Pollak, H.O. 1961 "Prolate spheroidal wave functions, Fourier analysis and uncertainty - I," Bell Syst. Tech. J. 40, 43.

Smith, H.A. 1966 "Improvement of the resolution of a linear scanning device," Siam. J. Appl. Math. 14, 231.

Smith, P.R., Peters, T.M., Bates, R.H.T. 1973 "Image reconstruction from finite numbers of projections," J. Phys. A 6, 361.

Sondhi, M.M. 1972 "Image restoration: the removal of spatially invariant degradations," Proc. IEEE 60, 842.

Stockham, T.G. Jr. 1972 "Image processing in the context of a visual model," Proc. IEEE 60, 828.

Stockham, T.G., Cannon, T.M., Ingebretson, R.B. 1975 "Blind deconvolution through digital signal processing," Proc. IEEE 63, 678.

Stroke, G.W. 1972 "Optical Computing," IEEE Spectrum 9(12), 23.

Stroke, G.W., Halioua, M. 1973b "Image improvement in high resolution electron microscopy with coherent illumination (low-contrast objects) using holographic image-deblurring deconvolution III B," Optik 37, 249.

Stroke, G.W., Halioua, M. 1973a "Image improvement in high-resolution electron microscopy with coherent illumination (low-contrast objects) using holographic image-deblurring deconvolution, III, part A, Theory," Optik 37, 192.

Stuller, J.A. 1972 "An algebraic approach to image restoration filter design," Computer Graphics and Image Processing 1, 107.

Tescher, A.G., Andrews, H.C. 1972 "Data compression and enhancement of sampled images," Appl. Opt. 11, 919.

Twomey, S. 1963 "On the numerical solution of Fredholm integral equations of the first kind by the inversion of the linear system produced by quadrature," J. ACM 10, 97.

Vander Lugt, A. 1968 "A review of optical data processing techniques," Opt. Acta 15, 1.

Watson, G.N. 1966 "A treatise on the theory of Bessel functions," Cambridge University Press, 2nd edn.

Zimmermann, F.S., Gupta, S.C. 1973 "A state variable approach to digital image processing," Computers and Electrical Engineering 1, 255.